



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 190532

**TO: Rosanne Kosson**  
**Art Unit: 1653**  
**Location: REM-3B84&3C18**  
**Serial Number: 09/856723**

**Tuesday, May 23, 2006**

**From: Beverly Shears**  
**Location: Biotech-Chem Library**  
**REM 1A54**  
**Phone: 571-272-2528**  
**beverly.shears@uspto.gov**

### Search Notes

Your queries have completed processing. You may access an electronic version via eDAN (SCORE) and /or <http://es/ScoreAccessWeb>. If the result files have been separated into two (2) or more versions, you may view additional files via the select "[View version list for this application](#)" link.

#### Protein Sequence Searches – February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (uniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

#### Published Applications Database - November 2005

Published\_Applications Nucleic Acid and Published\_Applications Amino Acid database searches now generate two sets of results each. The Published\_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications\_New databases; older published applications make up the Published\_Applications\_Main databases.

Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New).

Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.rapbm** (Published\_Applications\_AA\_Main) and **.rapbn** (Published\_Applications\_AA\_New).



**THIS PAGE BLANK (USPTO)**

STIC Biotech/ChemLib

190532

From: Kosson, Rosanne  
Sent: Thursday, May 18, 2006 3:55 PM  
To: STIC-Biotech/ChemLib  
Subject: request for sequence search- 09/856,723- REGULATORY PROTEIN PKE#83 FROM HUMAN KERATINOCYTES

Please search SEQ ID NO: 8 for me in the commercial and interference databases. This is a protein of 1076 amino acids. My mailbox is in 3C70. Thanks!!!

Rosanne Kosson  
Patent Examiner, AU 1653  
REM 3B84  
571-272-2923  
rosanne.kosson@uspto.gov

803402

CRFE

aa 1076

PD = 11/19/99 PCA  
12/7/98 DE ? no  
11/26/98 DE ? no

PCA/DE 99/03732

RECEIVED  
MAY 18 2006  
STIC

ME

5/12  
JH

\*\*\*\*\*

Searcher: \_\_\_\_\_  
Searcher Phone: \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date completed: \_\_\_\_\_  
Searcher Prep Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search  
NA# \_\_\_\_\_ AA# \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure #: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (Specify): \_\_\_\_\_

Date completed: \_\_\_\_\_

Searcher: Beverly e 2528

Terminal time: \_\_\_\_\_

Elapsed time: \_\_\_\_\_

CPU time: \_\_\_\_\_

Total time: \_\_\_\_\_

Number of Searches: \_\_\_\_\_

Number of Databases: \_\_\_\_\_

#### Search Site

\_\_\_\_\_ STIC

\_\_\_\_\_ CM-1

\_\_\_\_\_ Pre-S

#### Type of Search

\_\_\_\_\_ N.A. Sequence

\_\_\_\_\_ A.A. Sequence

\_\_\_\_\_ Structure

\_\_\_\_\_ Bibliographic

#### Vendors

\_\_\_\_\_ IG

\_\_\_\_\_ STN

\_\_\_\_\_ Dialog

\_\_\_\_\_ APS

\_\_\_\_\_ Geninfo

\_\_\_\_\_ SDC

\_\_\_\_\_ DARC/Questel

\_\_\_\_\_ Other CGN

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using BW model

Run on: May 20, 2006, 17:05:42 ; Search time 320 Seconds

(without alignments)  
3110.365 Million cell updates/sec

Title: US-09-856-723A-8

Sequence: 1 MKQYASPMPTQTDVTKFKP.....RTLEONKGMKAKKEKCVLQ 1076

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: UniProt 7.2.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5486	99.9	1196	2	Q6QNV2_HUMAN
2	5442.5	99.2	1231	2	Q6ND11_HUMAN
3	5265	95.9	1160	2	Q6N372_HUMAN
4	4919	89.6	962	2	Q69J77_HUMAN
5	4683.5	85.3	1242	2	Q692W3_MOUSE
6	4552	82.9	1206	2	Q6SOK3_MOUSE
7	4532	82.6	1206	2	Q91ZJ6_MOUSE
8	4414.5	80.4	893	2	Q51TV6_HUMAN
9	2255.5	41.3	1036	2	Q4SGB9_TRTNG
10	1154.5	21.0	762	2	Q5PQW3_FAT
11	1151.5	21.0	874	2	Q99J97_MOUSE
12	1131	20.6	759	2	Q99MS6_MOUSE
13	1127	20.5	771	2	Q99MS5_MOUSE
14	1082	19.7	723	2	Q8C2W6_MOUSE
15	1076.5	19.6	987	2	Q9V7X1_DROME
16	1075.5	19.6	987	2	Q6MR33_DROME
17	1018.5	18.6	987	2	Q7PKB6_ANOGA
18	966	17.6	310	2	Q9MW19_HUMAN
19	950.5	17.3	1716	2	Q99MS7_MOUSE
20	945.5	17.2	1415	2	Q9H7M7_HUMAN
21	941.5	17.2	1510	2	Q8N3D4_HUMAN
22	895	16.3	308	2	Q5SOK2_MOUSE
23	850.5	15.5	1919	2	Q4RIP0_TRTNG
24	627	11.4	125	2	Q53T97_HUMAN
25	579.5	10.6	288	2	Q4SGC0_TRTNG
26	556	10.1	243	2	Q8TB89_HUMAN
27	507.5	9.2	1010	2	Q9VU34_DROME
28	498.5	9.1	863	1	MILK1_HUMAN
29	480.5	8.8	967	1	Q6GN09_XENLA
30	480	8.7	870	1	MILK1_MOUSE
31	467.5	8.5	1886	2	Q47H5_TRTNG

32	453.5	8.3	1190	2	Q519A0_ENTH1
33	452	8.2	211	2	Q580X2_HUMAN
34	434	7.9	946	2	Q60W15_CABER
35	431.5	7.9	904	1	MILK2_HUMAN
36	418.5	7.6	1009	2	Q3UB65_MOUSE
37	416.5	7.6	1009	2	Q3TN34_MOUSE
38	414.5	7.6	1009	2	Q3TPY3_MOUSE
39	402.5	7.3	901	2	Q18749_CABER
40	402	7.3	583	2	Q4SMU9_TRTNG
41	389.5	7.1	992	2	Q5DUC2_MOUSE
42	379.5	6.9	242	2	Q4T0T5_TRTNG
43	368.5	6.7	989	2	Q94909_HUMAN
44	354.5	6.5	1738	2	Q76329_DICD1
45	344	6.3	2206	2	Q4REQ4_TRTNG

## ALIGNMENTS

Query Match	Score	DB 2	Length	1196	Matches	1075	Conservative	1	Mismatches	0	Indels	0	Gaps	0
1	5486	99.9%	1196	2	Q6QNV2_HUMAN	PRELIMINARY; PRT; 1196 AA.								
2	5442.5	99.2%	1231	2	Q6ND11_HUMAN	05-JUL-2004, integrated into UniProtKB/TrEMBL.								
3	5265	95.9%	1160	2	Q6N372_HUMAN	05-JUL-2004, sequence version 1.								
4	4919	89.6%	962	2	Q69J77_HUMAN	07-FEB-2006, entry version 9.								
5	4683.5	85.3%	1242	2	Q692W3_MOUSE	BH domain binding protein 1.								
6	4552	82.9%	1206	2	Q6SOK3_MOUSE	Name=BHBP1;								
7	4532	82.6%	1206	2	Q91ZJ6_MOUSE	Q51TV6_HUMAN								
8	4414.5	80.4%	893	2	Q51TV6_HUMAN	Q51TV6_HUMAN								
9	2255.5	41.3%	1036	2	Q4SGB9_TRTNG	Q4SGB9_TRTNG								
10	1154.5	21.0%	762	2	Q5PQW3_FAT	Q5PQW3_FAT								
11	1151.5	21.0%	874	2	Q99J97_MOUSE	Q99J97_MOUSE								
12	1131	20.6%	759	2	Q99MS6_MOUSE	Q99MS6_MOUSE								
13	1127	20.5%	771	2	Q99MS5_MOUSE	Q99MS5_MOUSE								
14	1082	19.7%	723	2	Q8C2W6_MOUSE	Q8C2W6_MOUSE								
15	1076.5	19.6%	987	2	Q9V7X1_DROME	Q9V7X1_DROME								
16	1075.5	19.6%	987	2	Q6MR33_DROME	Q6MR33_DROME								
17	1018.5	18.6%	987	2	Q7PKB6_ANOGA	Q7PKB6_ANOGA								
18	966	17.6%	310	2	Q9MW19_HUMAN	Q9MW19_HUMAN								
19	950.5	17.3%	1716	2	Q99MS7_MOUSE	Q99MS7_MOUSE								
20	945.5	17.2%	1415	2	Q9H7M7_HUMAN	Q9H7M7_HUMAN								
21	941.5	17.2%	1510	2	Q8N3D4_HUMAN	Q8N3D4_HUMAN								
22	895	16.3%	308	2	Q5SOK2_MOUSE	Q5SOK2_MOUSE								
23	850.5	15.5%	1919	2	Q4RIP0_TRTNG	Q4RIP0_TRTNG								
24	627	11.4%	125	2	Q53T97_HUMAN	Q53T97_HUMAN								
25	579.5	10.6%	288	2	Q4SGC0_TRTNG	Q4SGC0_TRTNG								
26	556	10.1%	243	2	Q8TB89_HUMAN	Q8TB89_HUMAN								
27	507.5	9.2%	1010	2	Q9VU34_DROME	Q9VU34_DROME								
28	498.5	9.1%	863	1	MILK1_HUMAN	MILK1_HUMAN								
29	480.5	8.8%	967	1	Q6GN09_XENLA	Q6GN09_XENLA								
30	480	8.7%	870	1	MILK1_MOUSE	MILK1_MOUSE								
31	467.5	8.5%	1886	2	Q47H5_TRTNG	Q47H5_TRTNG								

```

Db 361 RYKAKAPAPVLSKPTGLNENTVSAGKDLSTSPKSPSPAPVLAGRKPNASQSLVWCKE 420
Qy 301 VTKNRYGVKLTNTFTTSWRNGSLSCAILLHFRPDLIDYSLNPOLIKENKKAAYGFGASIG 360
Db 421 VTKNRYGVKLTNTFTTSWRNGSLSCAILLHFRPDLIDYSLNPOLIKENKKAAYGFGASIG 480
Qy 361 ISRLLESDMVLALIPDLTYMTLYLQIRAHFSGOELNVVQIIEENSSKSTKYGVNETDT 420
Db 481 ISRLLESDMVLALIPDLTYMTLYLQIRAHFSGOELNVVQIIEENSSKSTKYGVNETDT 540
Qy 421 NSVDQEKFAELSDLKREPELQOPISGAVPLISODSVFVNDGSGVSESEHQTDPDHL 480
Db 541 NSVDQEKFAELSDLKREPELQOPISGAVPLISODSVFVNDGSGVSESEHQTDPDHL 600
Qy 481 SPSTASPYCRRTKSDTEPQKSGQSSGRTSGSDPGICSNSTDTQAQVLLGKRLKAEYL 540
Db 601 SPSTASPYCRRTKSDTEPQKSGQSSGRTSGSDPGICSNSTDTQAQVLLGKRLKAEYL 660
Qy 541 ELSDLYVSDKKDKMSPPICETDEOKLQTLIDISNLEKEKLENSRSLKCRSDPESPICK 600
Db 661 ELSDLYVSDKKDKMSPPICETDEOKLQTLIDISNLEKEKLENSRSLKCRSDPESPICK 720
Qy 601 TSLSPTSKLGYSYRDLIDLAKKKKASLRQTESDPDADRTTLNHADHSSKIYQHRLLSROE 660
Db 721 TSLSPTSKLGYSYRDLIDLAKKKKASLRQTESDPDADRTTLNHADHSSKIYQHRLLSROE 780
Qy 661 ELKERAVLLLEQARDAALAKGNHNTATAPFCNROLSDQODERRROLERAROLIAE 720
Db 781 ELKERAVLLLEQARDAALAKGNHNTATAPFCNROLSDQODERRROLERAROLIAE 840
Qy 721 ARSGVKMSLEPSYGEAAEKLKERSKASGDBENDIETDNEIPIEGFVVGSGDELTYLEN 780
Db 841 ARSGVKMSLEPSYGEAAEKLKERSKASGDBENDIETDNEIPIEGFVVGSGDELTYLEN 900
Qy 781 DLDTPEQNSKLVLDKLEKLEVPQVANSPPSAQAQVTESSBQDMKSGTEDLATERLQK 840
Db 901 DLDTPEQNSKLVLDKLEKLEVPQVANSPPSAQAQVTESSBQDMKSGTEDLATERLQK 960
Qy 841 TTRFRFPVVPVSKOSTYRKTQLOSFQYINRPMKQORSIOETTKGNEKKAITETOR 900
Db 961 TTRFRFPVVPVSKOSTYRKTQLOSFQYINRPMKQORSIOETTKGNEKKAITETOR 1020
Qy 901 KPSEDEVLANGFQKOTSGYVVGELALLENQOKIDTRAAVLEKRLRYLMDGRNTESEAM 960
Db 1021 KPSEDEVLANGFQKOTSGYVVGELALLENQOKIDTRAAVLEKRLRYLMDGRNTESEAM 1080
Qy 961 MOEWFMVLNKKNALIRMNQSLLEKEHDLERRYEELNRELRALALIEDMQTEAQKRE 1020
Db 1081 MOEWFMVLNKKNALIRMNQSLLEKEHDLERRYEELNRELRALALIEDMQTEAQKRE 1140
Qy 1021 QLLLDLVALVNRKDALVRDLDAQEKQAEDEDEHLETTLEONKKNKAKKEKCYLQ 1076
Db 1141 QLLLDLVALVNRKDALVRDLDAQEKQAEDEDEHLETTLEONKKNKAKKEKCYLQ 1196

```

## RESULT 2

```

O8ND11_HUMAN PRELIMINARY; PRT; 1231 AA.
AC O8ND11;
DT 01-OCT-2002, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 17.
DE Hypothetical protein DKFZ566N1646 (NPR/calponin-like protein).
GN Name=EHBP1; Synonym=DKFZ566N1646, NACSIN;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;

```

```

RG The German cDNA Consortium;
RA Oettermajer B., Oettermajer B., Deuschlebauer S., Schajp A.,
RA Mewes H.W., Weil B., Amid C., Oeinger A., Fodor G., Han M., Wiemann S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Castellano-Munoz M., Fernandez-Chacon R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonCommercial 4.0 International license
CC
CC EMBL: AL833968; CAD38814.1; -; mRNA.
CC EMBL: AY31186; AAQ9741.1; -; mRNA.
CC HSP: O01082; 1BKX.
CC Ensembl: ENSG00000115504; Homo sapiens.
CC HNC: HNC:29144; EHBP1.
CC InterPro: IPR001715; Calponin_act_bd.
CC Pfam: PF00307; CH; 1.
CC SMART: SM00033; CH; 1.
CC PROSITE: PSS0021; CH; 1.
CC Hypothetical protein.
SQ SEQUENCE 1231 AA; 139658 MW; DB7A7DBE672AEB98 CRC64;

Query Match 99.2%; Score 5442.5; DB 2; Length 1231;
Best Local Similarity 96.6%; Pred. No. 4,1e-205;
Matches 1073; Conservative 1; Mismatches 2; Indels 35; Gaps 1;

Qy 1 MKQYASPMPTQTDVKKFKPLSKVVSALQFSLSCIFLREGKATBDMQSLASVSMKQ 60
Db 121 MKQYASPMPTQTDVKKFKPLSKVVSALQFSLSCIFLREGKATBDMQSLASVSMKQ 180
Qy 61 ADICNLDDPEEDNEDDENRVNQEERAKIT----- 91
Db 181 ADICNLDDPEEDNEDDENRVNQEERAKITGVINQNLMSLDEDDQDCIKANRSK 240
Qy 92 -----ELINKLNFLEAEKDLATVNSNPDDDAELNPFQDPDSEEPITETASPRKTE 145
Db 241 SASSSEELINKLNFLEAEKDLATVNSNPDDDAELNPFQDPDSEEPITETASPRKTE 300
Qy 146 DSFYNNSYNPFKEVQTPQYLNPPDEBAFVTIKDSPQSTKRNIRPVMSKYLVDSSK 205
Db 301 DSFYNNSYNPFKEVQTPQYLNPPDEBAFVTIKDSPQSTKRNIRPVMSKYLVDSSK 360
Qy 206 TEREELDESNPFPKSTPPNNLVNVOELTEREVKKAAPAPVLSKPTGLNENTVS 265
Db 361 TEREELDESNPFPKSTPPNNLVNVOELTEREVKKAAPAPVLSKPTGLNENTVS 420
Qy 266 AGCLSTSPKSPIPSPVLAGRKPNASQSLVWCKEYTKNRYGVKLTNTFTTSWRNGSLPCA 325
Db 421 AGCLSTSPKSPIPSPVLAGRKPNASQSLVWCKEYTKNRYGVKLTNTFTTSWRNGSLPCA 480
Qy 326 ILHFRPDLIDYKSLNPOLIKENKKAAYGFGASIGISRLLEPSDVLALIPDLTYMTYL 385
Db 481 ILHFRPDLIDYKSLNPOLIKENKKAAYGFGASIGISRLLEPSDVLALIPDLTYMTYL 540
Qy 386 YQIRAHFSGOELNVVQIIEENSSKSTKYGVNETDTNNSVDQEKFAELSDLKREPELQOP 445
Db 541 YQIRAHFSGOELNVVQIIEENSSKSTKYGVNETDTNNSVDQEKFAELSDLKREPELQOP 600
Qy 446 ISGAVPLISODSVFVNDGSGVSESEHQTDPDHLSPSTASPYCRRTKSDTEPQKSGQSS 505
Db 601 ISGAVPLISODSVFVNDGSGVSESEHQTDPDHLSPSTASPYCRRTKSDTEPQKSGQSS 660
Qy 506 GRTSGSDPGICSNSTDTQAQVLLGKRLKAEYLSDLYVSDKKDKMSPPICETDE 565
Db 661 GRTSGSDPGICSNSTDTQAQVLLGKRLKAEYLSDLYVSDKKDKMSPPICETDE 720
Qy 566 OKLQTLIDISNLEKEKLENSRSLKCRSDPESPICKTSLSPTSKLGYSYRDLIDLAKKKA 625
Db 721 OKLQTLIDISNLEKEKLENSRSLKCRSDPESPICKTSLSPTSKLGYSYRDLIDLAKKKA 780
Qy 626 SLRQTESDPDADRTTLNHADHSSKIYQHRLLSROEELKERAVLLLEQARDAALAKGNH 685

```

Db 781 SLRQTESPPDADRTLTNADHSSKIVQRLRLSROBELERARVLLLEQARRDALAKGNKH 840  
Qy 686 NTNTATPCNROLSPQOQDEERRRQULREARQULIABSGVMSLPSYGEAAEELKERS 745  
Db 841 NTNTATPCNROLSPQOQDEERRRQULREARQULIABSGVMSLPSYGEAAEELKERS 900  
Qy 746 KASGDENDNIDITNEIEIPGFFVVGDELNTLENDLDTTPONSRLVYLKCLLEVOPO 805  
Db 901 KASGDENDNIDITNEIEIPGFFVVGDELNTLENDLDTTPONSRLVYLKCLLEVOPO 960  
Qy 806 VANSPPSAQAQAVTSSSQDMKSGTEDLRTERLQKTERFRNPVVSFSDSTVRKQLOQSF 865  
Db 961 VANSPPSAQAQAVTSSSQDMKSGTEDLRTERLQKTERFRNPVVSFSDSTVRKQLOQSF 1020  
Qy 866 SOYIENPREMKORSIOBDTKKGNBEKKAITETQKPSBEDVANKGFPDTQYVVGELAA 925  
Db 1021 SOYIENPREMKORSIOBDTKKGNBEKKAITETQKPSBEDVANKGFPDTQYVVGELAA 1080  
Qy 926 LENEQOQIDTAAALVEKRLRYLMDTGRNTEEBEAMQEMFLVNGKALIRRNQOLSLLE 985  
Db 1081 LENEQOQIDTAAALVEKRLRYLMDTGRNTEEBEAMQEMFLVNGKALIRRNQOLSLLE 1140  
Qy 986 KEHDLERRRYELNLELRAMLAIEDWQKTBAQKREQLLDELVALVNRDALVVDLDAQE 1045  
Db 1141 KEHDLERRRYELNLELRAMLAIEDWQKTBAQKREQLLDELVALVNRDALVVDLDAQE 1200  
Qy 1046 KOABEBDEHLERTLEKONGKMAKKEKCVLQ 1076  
Db 1201 KOABEBDEHLERTLEKONGKMAKKEKCVLQ 1231

RESULT 3  
06NK72\_HUMAN  
ID 06NK72\_HUMAN PRELIMINARY; PRT; 1160 AA.  
AC 06NK72;  
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.  
DT 05-JUL-2004, sequence version 1.  
DT 07-FEB-2006, entry version 13.  
DE EMBP1 protein.  
GN Name=EBP1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;  
OC Homo.  
NCBI TaxID=9606;  
RN NCBI TaxID=9606;  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Uterus;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Strepiton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carrinci F., Prange C.,  
RA Rana S.S., Loquellano N.A., Peters G.J., Adameo R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Uterus;

RA Director MGC Project;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NonDerivs License  
CC -----  
DR EMBL: BC067215; AAH67215.1; -; mRNA.  
DR HSSP: Q01082; 1AA2.  
DR InterPro: IPR001715; Calponin\_act\_bd.  
DR Pfam: PF00307; CH; 1.  
DR SMART: SM00033; CH; 1.  
DR PROSITE: PS50021; CH; 1.  
SQ SEQUENCE 1160 AA; 132322 MW; DD79233FE0B14AF1 CRC64;  
Query Match 95.9%; Score 5266; DB 2; Length 1160;  
Best Local Similarity 96.6%; Pred. No. 3.2e-198;  
Matches 1039; Conservative 1; Mismatches 0; Indels 36; Gaps 1;  
Qy 1 MKQYASPMPTQTDVAKLKPFLSKKYVSAALQPSISCFILREGKATDEDMQSLASVWKQ 60  
Db 121 MKQYASPMPTQTDVAKLKPFLSKKYVSAALQPSISCFILREGKATDEDMQSLASVWKQ 180  
Qy 61 ADIGNLDDFEEDNEDDENRVNQBKAKITELINKLFLDEAKDLATVNSNPPDDPA 120  
Db 181 ADIGNLDDFEEDNEDDENRVNQBKAKITELINKLFLDEAKDLATVNSNPPDDPA 240  
Qy 121 AELNPPDPSSEETITETASPKRTEDSFYNSNYPFEVQTPQYLPNPPDEFAVTIKDS 180  
Db 241 AELNPPDPSSEETITETASPKRTEDSFYNSNYPFEVQTPQYLPNPPDEFAVTIKDS 300  
Qy 181 PPOSTKRNIRPVMSKYLYADSSKTEEBELDESNPFEKSTPPNNLVNPVOLETER 240  
Db 301 PPOSTKRNIRPVMSKYLYADSSKTEEBELDESNPFEKSTPPNNLVNPVOLETER 360  
Qy 241 RVKKKAPAPVLSPTKVTLMENVYASGKDLSTSPKSPSPSPVYGRKPNASQSLVWCKE 300  
Db 361 RVKKKAPAPVLSPTKVTLMENVYASGKDLSTSPKSPSPSPVYGRKPNASQSLVWCKE 420  
Qy 301 VTKKYRGVKTNTFTTSWRNGLSFCAIIHRRPDLIDVKSINLPQIKENNKAYGPGSIG 360  
Db 421 VTKKYRGVKTNTFTTSWRNGLSFCAIIHRRPDLIDVKSINLPQIKENNKAYGPGSIG 480  
Qy 361 ISRLBPSDMLVLAIPKLTVMYLYOIRAHFSGQELINVQIBNSKSTYKVGNYETD 420  
Db 481 ISRLBPSDMLVLAIPKLTVMYLYOIRAHFSGQELINVQIBNSKSTYKVGNYETD 540  
Qy 481 NSSYDQEKFYAELSDLRBPELQPIGAVDFLSQDDSVFVNSDGVESSEHQTPDHL 480  
Db 541 NSSYDQEKFYAELSDLRBPELQPIGAVDFLSQDDSVFVNSDGVESSEHQTPDHL 600  
Qy 481 SPSTASPYCRRYSKSTPEPKSQSSGRTSGSDDPGICSNSTDSQAQVYLGGKRLKAEYL 540  
Db 601 SPSTASPYCRRYSKSTPEPKSQSSGRTSGSDDPGICSNSTDSQAQVYLGGKRLKAEYL 660  
Qy 541 ELSLTVSDKKKMSPPFICBETDEQLOTLIDGSNIEKELTNSRLFCRSDSPSPKK 600  
Db 661 ELSLTVSDKKKMSPPFICBETDEQLOTLIDGSNIEKELTNSRLFCRSDSPSPKK 720  
Qy 601 TSLSPSTKLGYSYSRDLDAKKKASLRQTESPPDADRTLTNADHSSKIVQRLRLSROE 660  
Db 721 TSLSPSTKLGYSYSRDLDAKKKASLRQTESPPDADRTLTNADHSSKIVQRLRLSROE 780  
Qy 661 ELKERARVLLLEQARRDALAAGNKANTYATPPCNRQLSPQOQDEERRRQULREARQULIAE 720  
Db 781 ELKERARVLLLEQARRDALAAGNKANTYATPPCNRQLSPQOQDEERRRQULREARQULIAE 840  
Qy 721 ARSGVKKSELPSYGEAAEELKERSKASG----- 780  
Db 841 ARSGVKKSELPSYGEAAEELKERSKASG----- 865  
Qy 781 DLDTPQNSKLVDLKLKCLLEVOPOVANSPPSAQAQAVTSSSQDMKSGTEDLRTERLQK 840  
Db 870 -----EQNSKLVDLKLKCLLEVOPOVANSPPSAQAQAVTSSSQDMKSGTEDLRTERLQK 924

```

QY 841 TTERFRPVVFSKDSYVRKTQLOSPSOYIENRPEMKORSIOBDTKKGNBKAITETOR 900
DB 925 TTERFRPVVFSKDSYVRKTQLOSPSOYIENRPEMKORSIOBDTKKGNBKAITETOR 984
QY 901 KPSEDEVYANKGFQDTSQYVVGELAALENBQKQIDTRALVEKRLRYLMDTGRNTEEBEAM 960
DB 985 KPSEDEVYANKGFQDTSQYVVGELAALENBQKQIDTRALVEKRLRYLMDTGRNTEEBEAM 1044
QY 961 MOEWPMVNNKKNLIRRMNQSLLEKEHDLERRRELLNRELRAMLAIEDMQTEAQRRE 1020
DB 1045 MOEWPMVNNKKNLIRRMNQSLLEKEHDLERRRELLNRELRAMLAIEDMQTEAQRRE 1104
QY 1021 QLLDELVALVNRKDALVRDLDAQEKQAEDEDEHLEERTLEONKGMKKEBKVQ 1076
DB 1105 QLLDELVALVNRKDALVRDLDAQEKQAEDEDEHLEERTLEONKGMKKEBKVQ 1160

RESULT 4
094977 HUMAN PRELIMINARY; PRT; 962 AA.
AC 094977
DT 01-MAY-1999, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 24.
DE KIAA0903 protein (Fragment).
GN Name=KIAA0903;
OS Homo sapiens (human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA MEDLINE=93156230; PubMed=10048485; DOI=10.1093/dnares/5.6.355;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 5:355-364(1998).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC distributed under the Creative Commons Attribution-NonDerivative license
CC
DB EMBL; AB020710; BAA74926.1; -; mRNA.
DB HSSP; Q01082; 1BKX.
DB Ensembl; ENSG00000115504; Homo sapiens.
DB InterPro; IPR001715; Calponin_act_bd.
DB Pfam; PF00307; CH; 1.
DB SMART; SM00033; CH; 1.
DB PROSITE; PSS0021; CH; 1.
DB NON TER
SQ SEQUENCE 962 AA; 109413 MW; 588BC31SD322D7AE CRC64;

Query Match 89.6%; Score 4919; DB 2; Length 962;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 962; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DB 181 LVWCKEVTKNYRGVKTITNFTTSWRNGLSFCAILHFRPDLIDYKSLNPQDIKENNKAYD 240
QY 355 GFASIGISRLBESDMVLAIPDKLTWMTYLYOIRAHFGQELNVQIENSSKSTYKVG 414
DB 241 GFASIGISRLBESDMVLAIPDKLTWMTYLYOIRAHFGQELNVQIENSSKSTYKVG 300
QY 415 NYETDINSSVDQEKFAELSDLKREPELOPISGAVDFLSQDSDSVVNVDSGVESSEHQ 474
DB 301 NYETDINSSVDQEKFAELSDLKREPELOPISGAVDFLSQDSDSVVNVDSGVESSEHQ 360
QY 475 TPDDHLSPSTASPYCRRKSDTEPPQKSQSSGRTSGSDPDGICSNSTDSIOAYLLAKRL 534
DB 361 TPDDHLSPSTASPYCRRKSDTEPPQKSQSSGRTSGSDPDGICSNSTDSIOAYLLAKRL 420
QY 535 LKAEFTLESLDYSDKKQWSPPFCEETDEQQLDTLDGSNLEKRLERSLEGRSDP 594
DB 421 LKAEFTLESLDYSDKKQWSPPFCEETDEQQLDTLDGSNLEKRLERSLEGRSDP 480
QY 595 ESPFKKTSLSPTSKLGYSYRDLAKKQHASLRQTESPDADRTLLNADHSKIVQHR 654
DB 481 ESPFKKTSLSPTSKLGYSYRDLAKKQHASLRQTESPDADRTLLNADHSKIVQHR 540
QY 655 LLSRQSELKERAVLLLEQARRDALAKGNKNTNTATPCNRQLSDQDEERRRLREPA 714
DB 541 LLSRQSELKERAVLLLEQARRDALAKGNKNTNTATPCNRQLSDQDEERRRLREPA 600
QY 715 ROLIAARSGVKNSELPYGEAAEFLKERSKASGSENNIETDTEELPEGVVGGGE 774
DB 601 ROLIAARSGVKNSELPYGEAAEFLKERSKASGSENNIETDTEELPEGVVGGGE 660
QY 775 LTNLSENDLDTPEONSGLVDLKLKLEVOPOVANSPSSAOKAVTSSSEDDMKSGTEDLR 834
DB 661 LTNLSENDLDTPEONSGLVDLKLKLEVOPOVANSPSSAOKAVTSSSEDDMKSGTEDLR 720
QY 835 TEBRLQTTFRFRNPVVFSSKDSYVRKTQLOSPSOYIENRPEMKORSIOBDTKKGNBKA 894
DB 721 TEBRLQTTFRFRNPVVFSSKDSYVRKTQLOSPSOYIENRPEMKORSIOBDTKKGNBKA 780
QY 895 ITTQCKPSEDEVYANKGFQDTSQYVVGELAALENBQKQIDTRALVEKRLRYLMDTGRNT 954
DB 781 ITTQCKPSEDEVYANKGFQDTSQYVVGELAALENBQKQIDTRALVEKRLRYLMDTGRNT 840
QY 955 EEBEAMQEWPMVNNKKNLIRRMNQSLLEKEHDLERRRELLNRELRAMLAIEDMQTE 1014
DB 841 EEBEAMQEWPMVNNKKNLIRRMNQSLLEKEHDLERRRELLNRELRAMLAIEDMQTE 900
QY 1015 AQKRREQLLDELVALVNRKDALVRDLDAQEKQAEDEDEHLEERTLEONKGMKKEBKV 1074
DB 901 AQKRREQLLDELVALVNRKDALVRDLDAQEKQAEDEDEHLEERTLEONKGMKKEBKV 960
QY 1075 LQ 1076
DB 961 LQ 962

RESULT 5
0692M3 MOUSE PRELIMINARY; PRT; 1242 AA.
AC 0692M3;
DT 13-SEP-2004, integrated into UniProtKB/TrEMBL.
DT 13-SEP-2006, entry version 11.
DE KIAA0903 protein (Fragment).
GN Name=Shp1; Synonyms=mk1A0903;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Fetal brain;

```

RA Published:15368895; DOI=10.1093/dnares/11.3.205; Hiraoka S.,  
 RX Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,  
 RA Suga Y., Seino S., Nishimura M., Kaisho T., Hoshino K., Kitamura H.,  
 RA Nagase T., Ohara O., Koga H.;  
 RA "Prediction of the coding sequences of mouse homologues of KIA gene:  
 RT IV. The complete nucleotide sequences of 500 mouse KIA-homologous  
 RT cDNAs identified by screening of terminal sequences of cDNA clones  
 RT randomly sampled from size-fractionated libraries.",  
 RL DNA Res. 11:205-218(2004).  
 -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NonCommercial license  
 CC  
 CC EMBL: AK173055; BAD3233.1; -; mRNA.  
 DR Ensembl: ENSMUSG00000042302; Mus musculus.  
 DR WGI: MGI:267252; Embp1.  
 DR GO: GO:0005515; F:protein binding; IPI.  
 DR InterPro: IPR001715; Calponin\_act\_bd.  
 DR Pfam: PF00307; Ch.1.  
 DR SMART: SMO0033; Ch.1.  
 DR PROSITE: PS50021; Ch.1.  
 FT NON TER 1  
 SQ SEQUENCE 1242 AA; 140356 MW; 1C8D17FC63BC0932 CRC64;  
 -----  
 Query Match 85.3%; Score 4683.5; DB 2; Length 1242;  
 Best Local Similarity 84.1%; Pred. No. 2.2e-175;  
 Matches 945; Conservative 40; Mismatches 78; Indels 61; Gaps 10.  
 -----  
 QY 1 MKQYASPMPTQTDVLAKEKPLSKVVASALQPSLSCIFLRBEGKADDEDMQSLASIVSMKQ 60  
 DB 132 MKQYASPMPTQTDVLAKEKPLSKVVASALQPSLSCIFLRBEGKADDEDMQSLASIVSMKQ 191  
 QY 61 ADIGNLDDFEEDNEDDENRVNQEKAKIT----- 91  
 DB 192 ADIGNLDDFEEDNEDDENRVNQEKAKITETIVNQNALSLDEDDQDCIKQAVPSAK 251  
 QY 92 -----ELINLNLFLDEAKDLATYNSNPPDPDPAALNPPDPSEBPIITTSAPKRT 145  
 DB 252 SASSESELINLNLFLDEAKDLATYNSNPPDPDPAALNPPDPSEBPIITTSAPKRT 311  
 QY 146 DSFYNSNPNPKEVQTPQYLNLPFDEBPAFVITKDSPOSTRKXNIRPVMSSKYLVAADSSK 205  
 DB 312 ESFYNSNPNPKEVQTPQYLNLPFDEBPAFVITKDSPOSTRKXNIRPVMSSKYLVAADSSK 371  
 QY 206 TEEBELDESNPFEYBKSPPPNLNVNPOELETERVYRKA-----PAPVLS 253  
 DB 372 S-EEBELDESNPFEYBKSPPPNLNVNPOELETERVYRKA-----PAPVLS 429  
 QY 254 PKTGVNLENT-VSAGKDLSTSPKPSPIPSVLAGKPKNASOSILIVCKEYTKYRGVTKTN 312  
 DB 430 PKTGV-NENTVVASGKDLSTSPKPSPIPSVLAGKPKNASOSILIVCKEYTKYRGVTKTN 488  
 QY 313 FTSWRNLSPCALIHFRPDLIDVKSINPDIKKNKKKADGFAISIGISRLRESNVL 372  
 DB 489 FTSWRNLSPCALIHFRPDLIDVKSINPDIKKNKKKADGFAISIGISRLRESNVL 548  
 QY 373 LAIPDKLVTMTYLYQIRAFSGQELNVQIENSSKSYKGVNLETDTNNSVTDQEKFYAE 432  
 DB 549 LAIPDKLVTMTYLYQIRAFSGQELNVQIENSSKSYKGVNLETDTNNSVTDQEKFYAE 608  
 QY 433 LSDLKREBELQPISGAVDFLSGDDSVRVNDSGVGESSEHQTPDDHLSPSTASPYCRT 492  
 DB 609 LSDLKREBELQPISGAVDFLSGDDSVRVNDSGVGESSEHQTPDDHLSPSTASPYCRT 668  
 QY 493 KSDTPKQSQSGSGTSGDDPGICOSNDSQAOVLAKKKLLKRETELISLVYSDKKK 552  
 DB 669 KSDTPKQSQSGSGTSGDDPGICOSNDSQAOVLAKKKLLKRETELISLVYSDKKK 727  
 QY 553 DMSPEPFICEETDEBQKQLTIDIGSNLEKELIENSRELECRSDPESEPIKKTSHSPSTKLGYS 612  
 DB 728 DVSP-----LSAYBQGLQTVHAASDMEGCKMEKSNRELCRDLGELALITPNVSSPEKLG-- 781  
 QY 613 YSRDLDAKKGHASLRQTESDPDADRTLLNHAHSSKIVQHRLLSRQELKERARVAVLLEQ 672

[illegible]

```
Db 121 MKQYASMPQTQDVKLFKFKPLSKKVVSATLQFSLSLSCIFLREGKATDEDMQSLASLSMKQ 180
Qy 61 ADIGNLDDFEEDNEDDENRVNOBEKAKIT----- 91
Db 181 ADIGNLDDFEEDNEDDENRVNOBEKAKITELIVNQLNALSLSLDEDDDDCICQANVPSAK 240
Qy 92 -----ELINKNLFLEBAEKDLATVNSNPDDPAELNPFQDSEEPITETASPRKTB 145
Db 241 SASSEBELINTNLFLEBAEKDLATVNTNPFDEPVTETELNPFQDSE----- 287
Qy 146 DSFYNNSYNPFKEVQTYQYLNPFDEPAFTIKSPQOSTKTKIRIVDMSKITLYADSSK 205
Db 288 -----GVQTPQYLNPFDEPFTFVIMKDSPPQOSTRRKKLRPVDMSKITYADSSK 335
Qy 206 TEEBELDESNPFPKSTPPNNLVNVOLETERRVKKA-----PAPVLS 253
Db 336 S-BEELDESNPFPKPT-SPNNLVNTVQEGETERRVKRAAPAPAPAPAPAPALTP 393
Qy 254 PKTGVNENT-VSAGKDLSTSPKSPISPVLAGKRPNASQSLVWCKEYTKYRGVKTN 312
Db 394 PKTGV-NENTVVSAGKDLSTSPKSPISPVLAGKRPNASQSLVWCKEYTKYRGVKTN 452
Qy 313 FTTSMRNGSLFPCALHHRFPDLIDYKSLNPODIKENNKKAYDGPASIGISRLLEPSDMVL 372
Db 453 FTTSMRNGSLFPCALHHRFPDLIDYKSLNPODIKENNKKAYDGPASIGISRLLEPSDMVL 512
Qy 373 LAIPDKLTWMTYLYOIRAHFSGOELNVVOIENSSKSTYKGNVETDNTNSVDOEKVYAE 432
Db 513 LAIPDKLTWMTYLYOIRAHFSGOELNVVOIENSSKSTYKGNVETDNTNSVDOEKVYAE 572
Qy 433 LSDLKEPELQEPISGAVDELISQDDSVFVNDGSGESESHEQTPDDHLSPTASPYCRR 492
Db 573 LSDLKEPEPHQARAGAVDLISQDDSVFVNDGSGESESHEQTPDDHLSPTASPYCRR 632
Qy 493 KSTTEPKSQSSGRTSGSDPFCISNTDSTQAVLLGKRLKAEFTLISDLVSPDKK 552
Db 633 KSTTEPKSQSSGRTSGSDPFCISNTDSTQAVLLGKRLKAEFTLISDLVSPDKK 691
Qy 553 DMSPPFICEBTEDEQKOTLIGSNLEKEKLENSRSLCERSDPEPPIKSTLSPTSKLYGS 612
Db 692 DVSD---LSAYEOKLQTHVASSDMEGCKEKSRLSCRLDGBLAIYKPNVSSPSKLG-- 745
Qy 613 YSRDLDLAKKXKHAISLQTESDPADRTTLNADHSSKIVQHRLLSROBELKEBARVLLQ 672
Db 746 YNDDTPTKPKCASLRIQIESDPADKSTLNHADHPNKAQVQHRMLSRBELKEBARVLLQ 805
Qy 673 ARDDAALKAGNKINTNTATPFCNRQLSDQDEERRQLERARQLIAPASGVYMSLEPS 732
Db 806 ARDDAALKAGNKINTNTATPFCNRQLSDQDEERRQLERARQLIAPASGVYMSLEPS 865
Qy 733 YGMAAEKLEKERSKASDENDNIETDNEBIPEGFVVGDDDELTLNLEDDTDEBONSIV 792
Db 866 YGMAAEKLEKERSKASDENDNIETDNEBIPEGFVVGDDDELTLNLEDDTDEBONSIV 925
Qy 793 DLKILKLLLEVQPVANSPSSAAQKAVTESSEODMKSGTEDLARTERLOKTERFRNPVFS 852
Db 926 DLKILKLLLEVQPVANSPSSAAQKAVTESSEODMKSGTEDLARTERLOKTERFRNPVFS 985
Qy 853 KOSTVKTQLOSFQYIENRPEMKGRSIOEDTKGNEERAAITETQRPKSEDEVANKGF 912
Db 986 KOSTVKTQLOSFQYIENRPEMKGRSIOEDTKGNEERAAITETQRPKSEDEVANKGF 1042
Qy 913 KOSTVKTQLOSFQYIENRPEMKGRSIOEDTKGNEERAAITETQRPKSEDEVANKGF 972
Db 1043 KOSTVKTQLOSFQYIENRPEMKGRSIOEDTKGNEERAAITETQRPKSEDEVANKGF 1102
Qy 973 ALIRRNQSLSLKEHDLERREYELINRELAPMLAIEDMOKTEBAQKRREQLLDELVALVN 1032
Db 1103 ALIRRNQSLSLKEHDLERREYELINRELAPMLAIEDMOKTEBAQKRREQLLDELVALVN 1162
Qy 1033 KRDAVLDLDAOEKQAEDEHLETLQNKGNKAKKEKCVLQ 1076
```

```
Db 1163 KRDAVLDLDAOEKQAEDEHLETLQNKGNKAKKEKCVLQ 1206
RESULT 7
ID Q912J6_MOUSE PRELIMINARY; PRT; 1206 AA.
AC Q912J6;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 19.
DE KIAA0903-like protein.
GN Name=Bhdp1; Synonym=AF424697;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX MEDLINE=2325413; PubMed=12174196; DOI=10.1186/1471-2156-3-14;
RA Ruch S., Resch K., Thiel C., Ulbrich M., Platzer M., Jockusch H.,
RA Schmitz-John T.;
RT Comparative transcription map of the wobbler critical region on mouse
RT chromosome 11 and the homologous region on human chromosome 2p13-14.";
RL BMC Genet. 3:14-14(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF424697; AAL24806.1; -; mRNA.
DR HSSP; Q01082; 1BKR.
DR MGI; MGI:2667252; Bhdp1.
DR GO; GO:0005515; F:protein binding; IPI.
DR InterPro; IPR001715; Calponin_act_bd.
DR Pfam; PF00307; CH; 1.
DR SMART; SM00033; CH; 1.
DR PROSITE; PS50021; CH; 1.
SQ SEQUENCE 1206 AA; 136047 MW; 349A9A7C7CF0AC46 CRC64;

Query Match 82.6%; Score 4532; DB 2; Length 1206;
Best Local Similarity 82.0%; Pred. No. 1,8e-169;
Matches 922; Conservative 39; Mismatches 77; Indels 86; Gaps 11;

Qy 1 MKQYASMPQTQDVKLFKFKPLSKKVVSATLQFSLSLSCIFLREGKATDEDMQSLASLSMKQ 60
Db 121 MKQYASMPQTQDVKLFKFKPLSKKVVSATLQFSLSLSCIFLREGKATDEDMQSLASLSMKQ 180
Qy 61 ADIGNLDDFEEDNEDDENRVNOBEKAKIT----- 91
Db 181 ADIGNLDDFEEDNEDDENRVNOBEKAKITELIVNQLNALSLSLDEDDDDCICQANVPSAK 240
Qy 92 -----ELINKNLFLEBAEKDLATVNSNPDDPAELNPFQDSEEPITETASPRKTB 145
Db 241 SASSEBELINTNLFLEBAEKDLATVNTNPFDEPVTETELNPFQDSE----- 287
Qy 146 DSFYNNSYNPFKEVQTYQYLNPFDEPAFTIKSPQOSTKTKIRIVDMSKITLYADSSK 205
Db 288 -----GVQTPQYLNPFDEPFTFVIMKDSPPQOSTRRKKLRPVDMSKITYADSSK 335
Qy 206 TEEBELDESNPFPKSTPPNNLVNVOLETERRVKKA-----PAPVLS 253
Db 336 S-BEELDESNPFPKPT-SPNNLVNTVQEGETERRVKRAAPAPAPAPAPAPALTP 393
Qy 254 PKTGVNENT-VSAGKDLSTSPKSPISPVLAGKRPNASQSLVWCKEYTKYRGVKTN 312
Db 394 PKTGV-NENTVVSAGKDLSTSPKSPISPVLAGKRPNASQSLVWCKEYTKYRGVKTN 452
Qy 313 FTTSMRNGSLFPCALHHRFPDLIDYKSLNPODIKENNKKAYDGPASIGISRLLEPSDMVL 372
Db 453 FTTSMRNGSLFPCALHHRFPDLIDYKSLNPODIKENNKKAYDGPASIGISRLLEPSDMVL 512
Qy 373 LAIPDKLTWMTYLYOIRAHFSGOELNVVOIENSSKSTYKGNVETDNTNSVDOEKVYAE 432
```



```

Db      513 LAIPKLTVMYTYLYQIRAHFSGGELNVVQIENSSKSTYKGNVETDNTSSVDQKFFAE 572
Qy      423 LSDLRKEPELOQPIGAVDFLSQDDSVFVNDGVSSESEHQPDPDHLSPSPASRYCRR 492
Db      573 LSDLRKEPEPHOPANGAVDLISQDDSVFVNDGVSSESEHQPDPDHLSPSPASRYCRR 632
Qy      493 KSDTEPQKSGSGRTSGSDDPGICSNSTDSQAVLLGKKRLKAKETLSDLYVSDKKK 552
Db      633 KSDTEPQKSGSGRTSGSDDPGICSNSTDSQAVLLGKKRLKAKETLSDLYVSDKKK 691
Qy      553 DMSPEPFIEETDEQQLQTLIDIGSNLEKELKENSRLSECRSPESPIKTSLSPTSKGYS 612
Db      692 DVSP-----LSAYEQQLQTVHAASDMEQCKMEKRSLECRLLGELAITPKNVSSPSKLG-- 745
Qy      613 YSRDLDAKKGKASLRQTESDPPADRTLTNADHSKTVQHRLLSROELKERARAVLLEQ 672
Db      746 YNRDLDFPKPCASLRQTESDPPADRTLTNADHSKTVQHRLLSROELKERARAVLLEQ 805
Qy      673 ARDDAALKAGKGNHTNTATPPCNRLSDQDERRRQLREBARQLIAARSGVKMSLPS 732
Db      806 ARDDAALKAGKGNHTNTATPPCNRLSDQDERRRQLREBARQLIAARSGVKMSLPS 865
Qy      733 YGEMAAEKLKERSKASGDENDNIEIDTNEBIEGCVVGGDELITNLENDLTPGQNSKLV 792
Db      866 YGEMAAEKLKERSKASGDENDNIEIDTNEBIEGCVVGGDELITNLENDLTPGQNSKLV 925
Qy      793 DLKIKLLEVPQVANSFSSAAQKAVTSSSEODKMSGSEDITRTERLOQTTFRFPVTF 852
Db      926 DLKIKLLEVPQVANSFSSAAQKAVTSSSEODKMSGSEDITRTERLOQTTFRFPVTF 985
Qy      853 KSTVTRKTQLOSFQGYIENRPEMKRORSIOEDTKKGNKKAITETOKRPSDEVLNKG 912
Db      986 KSTVTRKTQLOSFQGYIENRPEMKRORSIOEDTKKGNKKAITETOKRPSDEVLNKG 1042
Qy      913 KDTSGYVVGELALLENQKQIDTRALVEKRLRYLMDTGRNTEBEAMQEMFMLVNNKN 972
Db      1043 KDTSGYVVGELALLENQKQIDTRALVEKRLRYLMDTGRNTEBEAMQEMFMLVNNKN 1102
Qy      973 ALIRRMQSLLEKSHDLERREYELNRLRAMLAIEDMQTBAQRRQLLDLVALVN 1032
Db      1103 ALIRRMQSLLEKSHDLERREYELNRLRAMLAIEDMQTBAQRRQLLDLVALVN 1162
Qy      1033 KRDLAVRDLDAQKQAEDEEHLSERTLEQNGKMAKKEKCVLQ 1076
Db      1163 KRDLAVRDLDAQKQAEDEEHLSERTLEQNGKMAKKEKCVLQ 1206

```

## RESULT 8

```

ID 053TV6_HUMAN PRELIMINARY; PRT; 893 AA.
AC Q53TV6;
DT 24-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 24-MAY-2005, sequence version 1.
DE 07-FEB-2006, entry version 3.
DE Hypothetical protein KIAA0903 (Fragment).
GN Name=KIAA0903;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Kalicki J., Johnson D., Harris A.;
RT "The sequence of Homo sapiens BAC clone RP11-443F16."
RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.H.;
RL Submitted (OCT-2000) to the EMBL/Genbank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.

```

```

RA Waterston R.;
RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Wilson R.K.;
RL Submitted (APR-2005) to the EMBL/Genbank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
DR EMBL; AC007098; AY14789.1; -; Genomic DNA.
DR Ensembl; ENSG00000115504; Homo sapiens.
DR InterPro; IPR001715; Calponin_act_bd.
DR Pfam; PF00307; CH; 1.
DR SMART; SM00033; CH; 1.
DR PROSITE; PS50021; CH; 1.
DR Hypothetical protein.
FT NON_TER 1
FT TER 893
SQ SEQUENCE 893 AA; 100446 MW; 88B387A63176A52E CRC64;
Query Match 80.4%; Score 4414.5; DB 2; Length 893;
Best Local Similarity 97.5%; Pred. No. 5.4e-165;
Matches 864; Conservative 6; Mismatches 15; Indels 1; Gaps 1;

```

```

Qy      66 LDDPEEDND-DDENRVNGEKKAKITELINKLFLDEAEKDLATVNSNPPDDDAEELN 124
Db      8 LSSIDEDQDDCTKQANMKSASASSEELINKLFLDEAEKDLATVNSNPPDDDAEELN 67
Qy      125 PFGDPEDEEPTTETASPRKTEDSFYNNSYNPFKEVQTPQYLNPPDEPEAFYTIKDSPPQS 184
Db      68 PFGDPEDEEPTTETASPRKTEDSFYNNSYNPFKEVQTPQYLNPPDEPEAFYTIKDSPPQS 127
Qy      185 TKRNIRPVDMSKYLVDSSKTEBEELDESNPFEYKSTPPNNLVNPFQELFETRRYKR 244
Db      128 TKRNIRPVDMSKYLVDSSKTEBEELDESNPFEYKSTPPNNLVNPFQELFETRRYKR 187
Qy      245 KAPAPVLSPTGVLNNTVYNSAGKDLSTSPKPSIIPSPVGRKYNASOSLLVWCKEYTKN 304
Db      188 KAPAPVLSPTGVLNNTVYNSAGKDLSTSPKPSIIPSPVGRKYNASOSLLVWCKEYTKN 247
Qy      305 YRGVKTINFTTWSNGSLFCAILHFRPDLIDYKSLNPQDIKENNKACAYDFASIGISRL 364
Db      248 YRGVKTINFTTWSNGSLFCAILHFRPDLIDYKSLNPQDIKENNKACAYDFASIGISRL 307
Qy      365 LEPSDVLALIPDKLTVMTYLYQIRAHFSGGELNVVQIENSSKSTYKGNVETDNTSSV 424
Db      308 LEPSDVLALIPDKLTVMTYLYQIRAHFSGGELNVVQIENSSKSTYKGNVETDNTSSV 367
Qy      425 DQKFFAELSDLRKEPELOQPIGAVDFLSQDDSVFVNDGVSSESEHQPDPDHLSPST 484
Db      368 DQKFFAELSDLRKEPELOQPIGAVDFLSQDDSVFVNDGVSSESEHQPDPDHLSPST 427
Qy      485 ASPYCRKTSKDTPEQKSGSGRTSGSDDPGICSNSTDSQAVLLGKKRLKAKETLSD 544
Db      428 ASPYCRKTSKDTPEQKSGSGRTSGSDDPGICSNSTDSQAVLLGKKRLKAKETLSD 487
Qy      545 LYSVDKKKMSPEPICEETDEQQLQTLIDIGSNLEKELKENSRLSECRSPESPIKTSLS 604
Db      488 LYSVDKKKMSPEPICEETDEQQLQTLIDIGSNLEKELKENSRLSECRSPESPIKTSLS 547
Qy      605 PTKSLGYSYSRDLDAKKGKASLRQTESDPPADRTLTNADHSKTVQHRLLSROELKE 664
Db      548 PTKSLGYSYSRDLDAKKGKASLRQTESDPPADRTLTNADHSKTVQHRLLSROELKE 607
Qy      665 RARVLEQARRDALKAGKGNHTNTATPPCNRLSDQDERRRQLREBARQLIAARSG 724
Db      608 RARVLEQARRDALKAGKGNHTNTATPPCNRLSDQDERRRQLREBARQLIAARSG 667
Qy      725 VKMSLEPSYGEMAAEKLEKRSKASGDENDNIEIDTNEBIEGCVVGGDELITNLENDLDT 784
Db      668 VKMSLEPSYGEMAAEKLEKRSKASGDENDNIEIDTNEBIEGCVVGGDELITNLENDLDT 727

```

```
QY 785 PEQNSKLVLDLKKLLEVOQOVANSPSSAOKAVTESSEODMSGTEDLRTERLQKTTTER 844
DB 728 PEQNSKLVLDLKKLLEVOQOVANSPSSAOKAVTESSEODMSGTEDLRTERLQKTTTER 787
QY 845 FRNPVFSKDSYVAKTQLOQSFQYIENRPEMKQRSIOEDTKGNEKAATETQKPS 904
DB 788 FRNPVFSKDSYVAKTQLOQSFQYIENRPEMKQRSIOEDTKGNEKAATETQKPS 847
QY 905 DEVLANKFQDTSQYVVGELALALENEKQIDTRALVKKRLRYLMDT 950
DB 848 DEVLANKFQDTSQYVVGELALALENEKQIDTRALVKKRLRYLMDT 893

RESULT 9
ID QASGB9_TETNG PRELIMINARY; PRT; 1036 AA.
AC QASGB9;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 1.
DE Chromosome 17 SCF14597, whole genome shotgun sequence.
GN ORFName=GSTEN0018711001;
OS Tetradon nigriviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
NCBI_TaxID=99883;
RN Nucleotide Sequence.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jallion O, Aury J.-M, Brunet F, Petit J.-L, Strange-Thomann N,
RA Mucelli E, Bouneau L, Fischer C, Ozouf-Costaz C, Bernot A,
RA Nicod S, Jaffe D, Fisher S, Lutfalla G, Doreat C, Segurens B,
RA Daillva C, Salanoubat M, Levy M, Boudet N, Castellano S,
RA Anthonard V, Ublin C, Castellani V, Katinka M, Vacherie B,
RA Blemont C, Skalli Z, Cattoiello L, Poulain J, De Berardinis V,
RA Cuend C, Duprat S, Broctier P, Couanceau J.-P, Gouzy J,
RA Parra G, Lardier G, Chapelle C, McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolf J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Lander V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigriviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN Nucleotide Sequence.
RP Genoscope; Whitehead Institute Centre for Genome Research;
RG Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
RL -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivative license
CC -----
CC EMBL, CAAB01014597; CAG00313.1; -; Genomic DNA.
CC InterPro; IPR001715; Calponin_act_bd.
DR Pfam; PF00307; CH; 1.
DR SMART; SM00033; CH; 1.
DR PROSITE; PSS0021; CH; 1.
SQ PROQUENCE 1036 AA; 114819 MW; C1A8E829F17262B6 CRC64;

Query Match 41.3%; Score 2265.5; DB 2; Length 1036;
Best Local Similarity 49.1%; Pred. No. 9.1e-81;
Matches 524; Conservative 121; Mismatches 188; Indels 235; Gaps 30;
```

```
DB 92 ADIGNLDDFEENBEBENNVNQEKAATITVILSLVCTFVALLEIKLNFLEDKQ 151
QY 106 DLN-TVNSNPDPDDAA----ELNPPDPPESEPIETETASPKRTEDSFYNSTNPPF--- 157
DB 152 EASFPAMSTNPFDEDDSLHNHFNPFEDPDMEABAPAFNSRQLDSDNDSNPFEPD 211
QY 158 --EVQPFQVNLNPFDEPEAFVTK--DSPPQSTRKQINRPVDMSKYLYADSSKTEBELDE 213
DB 212 EPELQTPG--NPFDEPDQDLDLPDEPEPKRQKGRVPVDMSKYLYADVNHNEDELD 269
QY 214 -----SNPFYPKSTPPNNLVNVDLETERRYKRAPAP 249
DB 270 FTQLPLGCVNSERSMNCFCILRSNPFYEPKTSFPAKPVGQ--PSLDMSSQKRAPP 328
QY 250 PVLSPKTG-----VLNNTYSA--GKUL-STSP-----K 275
DB 329 PSSSPGFPSPAPSKSSVDRQALAVGSPVTAIYGRBLASSPMTLQGIAVISLARK 368
QY 276 PSPSPVLGKRNASOSLLVCKEVTXNYRGYKTNFTTSMRNGLSFCAILHFPDLI 335
DB 389 PSPSPVLGKRNASOSLLAKCREVTXNYRGYKTNFTTSMRNGLAFCALLHFRDVI 448
QY 336 DYKSLNPQDIKENNKAYDGFASIGISRLLEPSDWLALPDKLTYWTYLYQIRAFPSGQ 395
DB 449 DYKSLNPQDIKENNKAYDGFASIGISRLLEPSDWLALPDKLTYWTYLYQIRAFPSGE 508
QY 396 ELNVVQIEENSSYSTYKVGNYETDTNNSVVDKFCYAEISLKKPELQGI---SGAVDF 452
DB 509 ELNVVQIEENSSYSTYKVGNYETDTNNSVVDKFCYAEISLKKPELQGI---SGAVDF 452
QY 453 LSQDSDV-----FVNSGVGESESEHQTPDDHLSPTASPCRTKSDTEPQSQSSGR 507
DB 569 TKNEDEGPKADCNVDKAVLTKTGHO----QSSPVPSP---RTGVYTSVLDSSQ--- 618
QY 508 TSGSDPFGICSNSTDSQOAVLLKRLKAELELSLDLYVSDKKKMSPFICEETDEQK 567
DB 619 TPASEERG-----SLTKANTLDELPERERE----- 645
QY 568 LQTLIDGSLNLEKLENSR---SLKCRSDPEPIKKTSLSP---SKLGYSRDLDA 620
DB 646 -----MEKROQKGVGEERTNGTEPSPGRGRLSPPHQKLGFSYNRRADLI 696
QY 621 KKKHSLRQTESPDPRDRT--TLNHADH-----SSKLVQHRLLSROBELKERAR 667
DB 697 KKKHSLRQTESPDPRDRT--TLNHADH-----SSKLVQHRLLSROBELKERAR 756
QY 668 VLEQARRDALAKGNKNTNTATPPCN--RQSDQDEERRQLRERARQLIAEARSQV 725
DB 757 LLEQARRDALAKGNKNTNTATPPCN--RQSDQDEERRQLRERARQLIAEARSQV 816
QY 726 KMELELSYGEAALAKLK-EKSKASGDEBNIEIDTMBELPBGVVGQDGLTYLENDLDT 784
DB 817 KMELELSYGEAALAKLK-EKSKASGDEBNIEIDTMBELPBGVVGQDGLTYLENDLDT 857
QY 785 PEQNSKLVLDLKKLLEVOQOVANSPSSAOKAVTESSEODMSGTEDLRTERLQKTTTER 844
DB 858 -----SERGIAGKRTAASHR----- 873
QY 845 FRNPVFSKDSYVAKTQLOQSFQYIENRPEMKQRSIOEDTKGNEKAATETQ-RKPS 903
DB 874 FRNPVFSKDSYVAKTQLOQSFQYIENRPEMKQRSIOEDTKGNEKAATETQ-RKPS 903
QY 904 EDEVLANKFQDTSQYVVGELALALENEKQIDTRALVKKRLRYLMDT 951
DB 934 EDEVLANKFQDTSQYVVGELALALENEKQIDTRALVKKRLRYLMDT 981

RESULT 10
ID QSPOM3_RAT PRELIMINARY; PRT; 762 AA.
AC QSPOM3;
DT 04-JAN-2005, integrated into UniProtKB/TrEMBL.
```



DT 04-JAN-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 12.  
 GN LOC0309169 protein.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Muridae; Murinae; Rattus.  
 NCBI\_Taxid=10116;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Heart;  
 RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marziani K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stopleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueding T.B., Toshiyuki S., Cantin P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butcherfield J.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Heart;  
 RG NIH MGC Project;  
 RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NonDerivative License  
 CC  
 CC EMBL; BC087115; AA087115.1; -; mRNA.  
 DR EMBL; ENSRNOG0000020788; Rattus norvegicus.  
 DR InterPro: IPR001715; Calponin\_act\_bd.  
 DR Pfam; PF00307; CH; 1.  
 DR SMART; SM00033; CH; 1.  
 DR PROSITE; PS00021; CH; 1.  
 DR SEQUENCE 762 AA; 84105 MW; 1249FD70B4B886C CRC64;  
 SQ  
 Query Match 21.0%; Score 1154.5; DB 2; Length 762;  
 Best Local Similarity 30.0%; Pred. No. 21e-37;  
 Matches 325; Conservative 104; Mismatches 205; Indels 449; Gaps 20;

QY 301 VTKAVGVKLTNTFTSMRNGLSFCAILHFRPDLIDYKSLNPDIKENNKKAYDGFASIG 360  
 DB 267 VTKRGVGTNTFTSMRNGLSFCAILHFRPDLIDYKSLNPDIKENNKKAYDGFASIG 326  
 QY 361 ISRLLEPBDVLLAI PDLVTWYLYLYRAHFSQGLNVVYIENNSKSTYKGVNFTD- 419  
 DB 327 VSRLLPBDVLLAI PDLVTWYLYLYRAHFSQGLNVVYIENNSKSTYKGVNFTD- 386  
 QY 420 ---TNSVSDQKFAELSDLKREP-ELQOPISGAVDPLSDSDSFVANDSGVGESESHQT 475  
 DB 387 PDCINAGDLAQRLEHGEASTKEKAVNRGTGAI PKVASRDTLGSSSKXGEAEVQEA 446  
 QY 476 PDLHSPSTASPYCRRTSTDPKQSSSGRTGSDPDGICSTNTDSTQAVLLGKRLL 535  
 DB 447 ---VQGEASDQPRASSTSP--VPTSELVNGVGPASVSGV----- 484  
 QY 536 KATLELSDLVVSDKDDMPPICEBTDEQKQLTLIGSNLEKKEKLENSRLCRSDPE 595  
 DB 485 -----RLRRS-SVNGEAGPV 498  
 QY 596 SPIKKTSLSPSKLGYISRDLDLAKKXASLRQTESDPPADRTTLNADHSKIVQRL 655  
 DB 499 PP-----PRAHGSFHVPRADILKTRRSRLRNSNS----- 528  
 QY 656 LSRQELKERARVLLQARNDALAKGNKNTNTATPCNRQLSDQDEBRRQLREBAR 715  
 DB 529 ----- 528  
 QY 716 QLIAEASGVMSSELPSYGEWAKEKLEKRSKASGDBNDNIEIDTNEBIPGFGVGGDEL 775  
 DB 529 -----FSVDDQDS-----CTAVGAGPAG 546  
 QY 776 TNLENDLTPBQNSKLYDLKIKLLVQPOVANSPPSAQAQAVTESSEQDKSGTEDLRT 835  
 DB 547 SGA---VEGWPDSLSLSTGLTATAPQPPSGNP----- 578  
 QY 836 ERLKTERFRANPVVFSKDSIVRTQLQSPSYIENRPEMKQRSIQEDTKGNEBKAI 895  
 DB 579 ---TESSSPN----- 586  
 QY 896 TETQRKSEDEVLNKGKDTISQYVVGELALLENKOQIDTRPALVKKRLTYMDTGRNTE 955  
 DB 587 -----GEAQQRQDTSQYVCAELQALBEGQSIDGAALAEVKKQLSLMESGNRL 638  
 QY 956 EEEANQGEFNLVKKKALIRRNQQLSLREKHDLEERYLLNRELPAALIEDWQTEA 1015  
 DB 639 QEBLLIQEWFTLVKKKALIRROPQLLIEBQDLERRFEILSRLEPAALIEBWQTEA 698  
 QY 1016 OKREBOILLDELVALVKKRDALVPDLDAQKQAEEDHELERTLEQKKGKAK---KEEK 1072  
 DB 699 QQRHEQLLELVSLVQRBDELVDQKERIALEBDBLERLGEQRRRKVSRLSRER 758  
 QY 1073 CVL 1075  
 DB 759 CTL 761  
 RESULT 11  
 Q99J97 MOUSE  
 ID Q99J97\_MOUSE PRELIMINARY; prt; 874 AA.  
 AC Q99J97  
 DT 01-JUN-2001, integrated into UniProtKB/TrEMBL.  
 DT 07-FEB-2006, entry version 1.  
 DE Tangerin B (LOC114601 protein).  
 GN Name=Kcnk7;  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Muridae; Murinae; Mus.  
 NCBI\_Taxid=10090;  
 RN [1]

RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=BALB/c;  
RA Melichar J.M., Noegel A.A., Korenbaum E.;  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;  
RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Kraemer R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marisano K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uadiri T.B., Toshlyuk S., Carrincci P., Prange C.,  
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Boak S.A., McBain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Gilwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;  
RA Strauberg R.;  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms  
CC Distributed under the Creative Commons Attribution-NonDerivs License  
CC  
DR EMBL; AF305088; AAK32110.1; -; mRNA.  
DR EMBL; BC004660; AAH04660.1; -; mRNA.  
DR HSSP; 001082; 1BKX.  
DR MGI; MGI:1341841; Kcnk7.  
DR GO; GO:0016021; C:Integral to membrane; RCA.  
DR GO; GO:0016020; C:membrane; RCA.  
DR GO; GO:0005267; F:Potassium channel activity; RCA.  
DR GO; GO:0006812; P:cation transport; RCA.  
DR GO; GO:0006813; P:potassium ion transport; RCA.  
DR InterPro; IPR001715; Calponin\_act\_ba.  
DR Pfam; PF00307; CH; 1.  
DR SMART; SM00033; CH; 1.  
DR PROSITE; PSS0021; CH; 1.  
SQ SEQUENCE 874 AA, 95561 MW, 9480358BCDF586D CRC64;  
Query Match 21.0%; Score 1151.5; DB 2; Length 874;  
Best Local Similarity 30.3%; Pred. No. 3.2e-37;  
Matches 337; Conservative 116; Mismatches 264; Indels 395; Gaps 27;  
QY 1 MKGYASMPPTQTDVTKLKFPLSKKVVSAALQFSLSCLFLREGKATDEDMQSLASVSMKQ 60  
DB 120 LAHAGGVPAQVLPKRLKPKSVVAHAEISLISGLVLRBGRATIDDMQSLASVSMKVP 179  
QY 61 ADIGNTLDPEEDNEDDENRNOEKA-AKITELINKLNFIDEAKDLATVNSNPPDDP 119  
DB 180 SDVGNLDDFAESBEEBANGAPAEVTRGQSDLSRLKTLCE-EBDEGHIR-----PQ 232  
QY 120 AAEINPFGDDP-SSEPTETASPKTSDSYNNSYNPFKEVQTPQVYNPDEPAVYTIK 178  
DB 233 QAARSSASADTSPAPVASAPAPVRA-----FRGQSSBPAALTGCG 273  
QY 179 DSEPGSTKRNKINRPVMSKVLVADSSKTBEEELDESFPYEPKTPPPNNLVNPVQELER 238  
DB 274 VGP-----ETPPEPPSP--ETSTGPGQTMVPTP--AP 304  
QY 239 ERVRKAKAPAPV-----LSPKTVLNNENTVSA-----GKDL-STSPKPSPIP 280

DB 305 RLKRGDAPSSPVPCSGDBVPTNSDEPPTGMSSGGERQAQISQEGTEAHEARPEPIEA 364  
QY 281 SPVLGRKRNASGSLVWCXVYTKYRGVKTNTTTSWRNGLSRCAIHHRRPLDIYKSL 340  
DB 365 PPTL---VSSQSILWCQCEVINGRGVCTITNPTTSWRNGLARCAIHHRRYPDKIDYFSL 421  
QY 341 NPDIKENNKAAVDGFASISIRSLRPSDMVLLAIPDKLTVMYLYVQIRAHFSGQELNV 400  
DB 422 DPLNITKQNKQADGPAALGVSLRLEPADVLLSVPKLIVMYTLLQIRAFCTGQELQV 481  
QY 401 QIEHNSKSTYKGVNVTDTNSSVDQEKFYAEISDLKREBELQOPT-----SGAVD 451  
DB 482 QLEGGGSGSYRVGNAPSLPDCLDA---GDIAQLRLEHGAIVPEPKAVNRGTGAI 537  
QY 452 FLSDQDSVFVNDGVGSESEHQTDPDHLSPSTASPCRTKSDTEPKQSGSSGRSSG 511  
DB 538 KVASRBDTSLSCSKRDGAEE-----VAQBAIPQBAPTDGPARRSS 576  
QY 512 DDP-----GICSTDTSTQAQVLLGKRLKAEITLISDLVVSXKCKMSPPICEETDEQ 566  
DB 577 TTIVVPAEGLVNGVGS-----GVRLKRS----- 601  
QY 567 KLTGLDIGNLKEKLENSRSLCERSDPESPIKTSLSPTSGLGYSRDLDAKKHAS 626  
DB 602 -----SYNGBAQVPP-----PRAHGSFHVRRADLLKRRSR 634  
QY 627 LRQTSDDPADRTTLTAHADHSKIVQHRLLSRBELKERARVLLQARRDALAKGNKH 686  
DB 635 LR----- 637  
QY 687 TTTATPFCNRQLSDODDEBRRRLRRARQLIAARSGVMSSELPSGVMAAEKLEKRSK 746  
DB 638 SNRPS-----VDDQD-----SGAAVAGPA-GPGAV- 663  
QY 747 ASGDENDNIEIDTNEBIPFGFVVGGBDELTNLENDLTPBQNSKLVDLKLLLEVQPV 806  
DB 664 ----- 667  
QY 807 ANPSSAAQAQVATESSEBQMKSGTEDLRTERLOKTERFRANPVVFSKOSTVTRKTQLQSF 866  
DB 668 ASSPDNPPLPAPVPOQPPGPPPT- 692  
QY 867 QYLENRPENKORSIEDTQKGEAKAITEQRKSPSEDEVLNKGFPTDSQYVVGELAL 926  
DB 693 ---ESSPSLGEFGDQ- 721  
QY 927 ENEQKIDTRALVERKRLRYLMDTGNTBEEBAMQEWFLVNKKUALIRRNQLSLK 986  
DB 722 EQGQGIQDRAAEVKEQLSLMESGANRLQEBVLQIEMFTLVNKKUALIRROQQLDLIE 781  
QY 987 EHLDERRYELNRELDAALAIEDMQTEQAQRREQLIDELVALVNKRDAVDDDAQEK 1046  
DB 782 EQLERRFELLSLELRLAMLAIEEMQTKVAQHRQELBELVSLVNQRDELVDLDQKER 841  
QY 1047 QAEEDHELERTLQONKGMKAK---KEBKV 1075  
DB 842 IALEEDERLERGLQRRKRVSKLSRRERCTL 873  
RESULT 12  
Q99MS6\_MOUSE PRELIMINARY; PRT; 759 AA.  
AC Q99MS6;  
DT 01-JUN-2001, integrated into UniProtKB/TrEMBL.  
DT 07-FEB-2006, entry version 17.  
DE Tangerin C.  
GN Name=Kcnk7;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Murinae; Mus.



Matches 327; Conservative 108; Mismatches 195; Indels 466; Gaps 26;

Qy	1	WKQYASPMPTQTDVYLKXKRPJLSKXYVSAALQSLCIPFLREKADJEDMOSJASIVSMKQ	60
Db	120	LAHHAQVPAQVPLRLKPKXSVKVYVHBLSLTSLGVLBRRAATDDMQJSLASIMSYVP	179
Qy	61	ADIGNLDDFEEDNEDDENRVNQEBAKATTELINKLFLDRAEKDLATVANSPPDDPA	120
Db	180	SDVGNLDDPFASDEB-----	194
Qy	121	AELNPFQDDPDEBEPITETASPKTDESPFNNSYNPFKEVQTOQYINPDPDEBAFTIIDS	180
Db	195	-EANGPGAPB-----	207
Qy	181	PPQSTK-----RKNIRPVMISKLYADSSKTEEBELDESINPFYKSTPPNNLVNVPBEL	236
Db	208	GPQSGRGCAPIRLGRPF-DLSREL-----KTLCEEBDECH-----IRPQ--	245
Qy	237	ETERRVKKAPAPVPVLSPKYGLNENTVYSAKDLSTSPKSPSIPSPVLGRKPNASQSLIV	296
Db	246	-----AAARP-----SSABD--TSPAPAP--PTL---VSSSQSLILE	274
Qy	297	WCKEYTKNYRGVKITNFTTSMRNGLSFPCALLHFRPDLIDVLSLNPDLIKENNKAYGCF	356
Db	275	WQEBYNGIRGVCITNFTTSMKNGSLAFCAILHFRPDKIDFSLDPLIKONNKAPFQGF	334
Qy	357	ASIGSRLLBPSDWLTLAIIPDLITVMTLYQIIRAHSGQGLNVLVYIIBENSSKSTYKENVY	416
Db	335	AALGVSRLLBPAWLVLSVPRDLIMYMTLYCQIRACTQGBELQVYLBEGGGSGTYRVGNA	394
Qy	417	ETDINSSVDQEKFYALBSDLKREPLQOP-----SCAVDLSODDSYFVNDSSVG	467
Db	395	QPSLPPDCIDA---GDLAQRUREHAEVPTBEKNAVNGTGAIPVYASRDPLDSCSSXDG	450
Qy	468	ESSESHQTPDHLSPSTASPYCRATKSDTEPQKSOSSGRTSGSDP-----GLCSINTDS	522
Db	451	EAE-----VAQBAIPQEAFTDPBARASSTTPVYPAEGLVNGVA	489
Qy	523	TOAOVLLKQKLLKAETLELSDLYVSDKKDMSPPFICEBTDQKQLQTLIDIGSNLEKEKL	582
Db	490	S-----GGVLRRS-----	498
Qy	583	ENSRLEBCRSDPESIKKTSLSPTSKLGYSYSRDLDLKKKQYASLROTESPDADRTTLN	642
Db	499	-----SYNGEAGVPP-----PRAHGSFHYKRDLLKRRSRK-----	533
Qy	643	HADHSSKIVOHRLSLROBELKERARVLLQOARRDALXAKNGKNTNTATPCPNROLSDQ	702
Db	534	-----NSNFS-----YDDQ	543
Qy	703	DEBRROQLREBARQOLIABRSQVMSKELPSYGEAAEKLKERSKSGCBENDIELDTNBE	762
Db	544	D-----SGAAAGAPR--GPGAVE-----	560
Qy	763	IPEGFVVGCGDELITNLENDLDTPEQNSKVLKDKGLLEVOQVANSPPSAAQKAVTESS	822
Db	561	-----GNPZSSSDANPLPAPVYQ	580
Qy	823	BQMKSGSTEDLRTERLQKTERFRNPVVPFSKDSYTRKQLOSFSQYIENRPMKRORSIQ	882
Db	581	PPGPPPPPE-----SSPSLGEAGLQ	602
Qy	883	EDTKKGNREKAAITETQKRPSEBDVLANGPOTQGYVVGELAALEBKQOITDRAALVEX	942
Db	603	R-----FOQTSQVCAELQALQEBQOQIDGRAAEVEX	634
Qy	943	RLRYLMDTGRNTEEBAMQEMFMLVNKKALLIRMNQLSLBEKHEDRERYELANLELR	1002
Db	635	QLRSLMESGARLQGEVLIQEMFTLVNKKMLIRQDQLOLILIEQDLERRPILLSREL	694
Qy	1003	AMLALIEDMOKTEAQGRQOLLIDELVALVNGKDALVRDLDAQEKABEDHLERTLEON	1062
Db	695	AMLALIEBQKVAOQHREOLLLEBELVSLVNRDELRVRLDQERLALIEBDRLEGRLEQOR	754

QY 1063 KGMKAK---KEKCVL 1075  
ID : : : : :  
Db 755 RRVSRQLSRRCCTL 770

RESULT 14  
Q8C2W6\_MOUSE  
AC Q8C2W6\_MOUSE PRELIMINARY; PRT; 723 AA.  
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.  
DT 01-MAR-2003, sequence version 1.  
DT 07-FEB-2006, entry version 16.  
DE 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched  
DE library, clone:R330024K12 product:cangerin, full insert sequence.  
DE (Fragment).  
GN Name=Kcnk7;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RN NCULEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Ovary;  
RX MEDLINE=99273253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RT Carninci P., Hayashizaki Y.,  
RL "High-efficiency full-length cDNA cloning",  
RN Methods Enzymol. 303:19-44(1999).  
RN [2]  
RN NCULEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Ovary;  
RX PubMed=16141072; DOI=10.1126/science.1112014;  
RN Carninci P., Kauckawa T., Katayama S., Gough J., Frith M.C., Medea N.,  
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,  
RA Bajic V.B., Bremner S.E., Batalov S., Forrest A.R., Zavolan M.,  
RA Davis M.J., Wilming L.G., Altshuler V., Allen J.E.,  
RA Amesbury-Imphombao A., Aweisler R., Acurayiya R.N., Bailey T.L.,  
RA Banasi M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,  
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,  
RA Crowe M.L., Dalla B., Dairymple B.F., de Bono B., Della Gatta G.,  
RA Di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulner G.,  
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,  
RA Geisler-Hemming P., Gingeras T.R., Gojobori T., Green R.B.,  
RA Guetlich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,  
RA Hill D., Hummel L., Iacono M., Ieko K., Iwama A., Ishikawa T.,  
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,  
RA Kitano H., Kollas L., Krishnan S.P., Kruger A., Kummerfeld S.K.,  
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,  
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,  
RA Matsuda H., Matsumura S., Miki H., Mignone F., Miyake S., Morris K.,  
RA Mottazari-Rad R., Mulder N., Nakano N., Nakanishi H., Ng P.,  
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohata O.,  
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,  
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,  
RA Roet B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,  
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,  
RA Shibata Y., Shimada H., Shinada K., Silva D., Sinclair B.,  
RA Sperling S., Stupka E., Sugizawa K., Sultana R., Takenaka Y., Taki K.,  
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,  
RA Ueda H.R., van Nimwegen B., Vezardo R., Wei C.L., Yagi K.,  
RA Yamashita H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,  
RA Grimmond S.M., Tesardale R.D., Liu E.T., Bruce V., Quackenbush J.,  
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,  
RA Fukuda S., Kanamori-Katsuyama M., Suzuki M., Aoki J., Arikawa T.,  
RA Iida Y., Iimura K., Itoh M., Kato T., Kawai I., Kawagashita N.,  
RA Kawashima T., Kojima M., Kondo S., Kono H., Nakano K., Ninomiya N.,  
RA Nishio T., Okada M., Plessey C., Shibata K., Shikano T., Suzuki S.,  
RA Tasami M., Waki K., Watanishi A., Okamura-Oho Y., Suzuki H., Kawai J.,  
RA Hayashizaki Y.;  
RT "The transcriptional landscape of the mammalian genome.",  
RL Science 309:1559-1563(2005).  
RN [3]







RA Yu C., Rubin G.;  
 RT "Drosophila melanogaster release 4 sequence."  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RL Submitted (JAN-2006) to the EMBL/Genbank/DBJ databases.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NonCommercial License  
 CC -----  
 CC EMBL; AE003804; AAF57918.2; -; Genomic\_DNA.  
 DR HSSP; Q01082; 1BKX.  
 DR FlyBase; FBgn0034180; CG15609.  
 DR InterPro; IPR001715; Calponin\_act\_bd.  
 DR Pfam; PF00307; CH; 1.  
 DR SMART; SM00033; CH; 1.  
 DR PROSITE; PS50021; CH; 1.  
 DR Chaperone.  
 KW SEQUENCE.  
 SQ 967 AA; 111231 MW; D1B7AF2053AD0CE CRC64;  
 Query Match 19.6%; Score 1076.5; DB 2; Length 987;  
 Best Local Similarity 28.2%; Pred. No. 3.2e-34;  
 Matches 324; Conservative 178; Mismatches 291; Indels 357; Gaps 38;  
 QY 1 MKQYASPMPTQDVYKLRKPKLSKKVVSAALOPGLSCIFLRGKADDEMOSLASIVSMKQ 60  
 DB 121 MKRTASISSTQSFSLSPKSKITTAASLELTISCVLRGKATDEDMOSVSMVSNN 180  
 QY 61 ADIGNLDDFER-----DNEEDDENRVNQ-----BEKAK 89  
 DB 181 NDVAPLDLELDLPDLGFSSEITDNDPTQQLNHTTSLNGCIDVATQSVPSLSEDP 240  
 QY 90 ITELINKLPLDEAKDIAVNSNPFDDPAEINPFCDPS-----EERI-- 135  
 DB 241 LAESFNPIHF--ELPADGKREANALDLPAAAGSSGABESLKTPTNGLQHVVDQTPKS 298  
 QY 136 -TETASPKTB-DSPYNNSYNPFKEVQTPQYLNPFDEBE-----AFVTIKDSPQS 184  
 DB 299 PDEVKQPTVESLRSTNEDFDKMTSTPLEPVNKDVKPKKRALPFTEKED-BOD 357  
 QY 185 TKRKAIRVDMSKYLYADSSKTEBEELDESNPFYEPKSTPPNNLVNPFQELETER--- 240  
 DB 358 LRVSVYK-DTTKTI-GDNSTTKKPKESQSSKQDASSVVPASARPELOPLANKKSYEP 415  
 QY 241 -RVKRAKAPPVLSPKTGVNLENTVSAGDLSTSPKS---PIPSVYGRKPNASGSL 295  
 DB 416 SOTPESAPVP-----ASVINE--SIGSCFSTGSLNNSLKPVEKIVL-KENTPGQDL 465  
 QY 296 VMCEKVTNKGKVTINTTTSMRNGLSRCALIHHRPDLIDYKSLNPQDICKENKKAADG 355  
 DB 466 EMCKEVTVDYVNVKTNLTTSWRNGMAFCALIHFPPLIDMSKLSADVVGNCRIGFDA 525  
 QY 356 PASIGISRLPEPDMVLAIPOKLTVMTYLYQIRAHFSGOELNVYQIENSSTKYKVN 415  
 DB 526 AESIGIPVIEPRDMNMLTVPDKLAVMTYLHQLRHAFGKQIKIQIGSTADESSYVIGD 585  
 QY 416 YETDNTSSVDQEKFYAELSDIKREPELOQPISGAVFLSODSVFVNDGVSSESEHOT 475  
 DB 586 YKSD-NLSQNRINLFHPSKSLHLHQNSPDEBING-----NKSQGLS 625  
 QY 476 PDDHLSPSTAPYCKRTKSDTEBPQKSSQSGRTSGSDPGICSNSTDSIOAOVLAKKRL 535  
 DB 626 P-----TSKQDVKNL-----ILVNSKNIL 644  
 QY 536 KAEFLSLDLYVSDKKQMSPPFICERTDQKLOTLDIGSNLBEKELNSRLSECRSDPE 595  
 DB 645 -----DKVLSP-----TKDKNSINASQHANQSPMLSCQT--- 673  
 QY 596 SPIKTSLSPTSKLGYSYRDLAKKKHSLRQTESDPPADRTTLNADHSSKIVQRL 655  
 DB 674 TPPODESLSDKGSPQ-----GKENSSTI--TTIDPQA-----ASRI 706

QY 656 LSRQELERARVILLQCARRDPAALKAGKNTNTATPFCNRLQSDQDEBRRRLRRAR 715  
 DB 707 LTRKMSKAKLMMERKALN-----KTSDANBERQGRLEQAR 746  
 QY 716 QLIAEARSQVMSLPSYGEAAEKLERSYASGDENDIBIDTNEBIPSGFVVGSDLE 775  
 DB 747 RLILETR-----VKSQG--- 758  
 QY 776 TNLENDLTPRQSKLVLDKLLLEVOPOVANSPPSAAQKAVTESSRODKSGTREDART 835  
 DB 759 -----SIESPTRP-----KLERTI-----SP-----IHNGABEFYA 784  
 QY 836 ERLQK-----TTERFRNPVFSKDSVTAKTQLQSPQYIENRPEMKRQRSIOEDTKK 887  
 DB 785 EKAKELEKESGSHSLSDPKV-----LQSNALVE-----RVSPKHKR 825  
 QY 888 GNEEKAATTEQRRKPSDEVLANKGFKOTSQYVGEALALENEQKQIDTPAALVEKRLRYL 947  
 DB 826 G-----DGLSYIDSELEALEREQEAIDQKASNLKAKRAV 860  
 QY 948 MDTGKNTRE-EEAMQEWPMLVNKNALIRRMQSLSEKHDLEKRYELLNRELRAMLA 1006  
 DB 861 MGNPKTEBTEBQLSQWFTLVNKNALLRRQWQNLTLQEKDLERKTYMLNOELRAAQ 920  
 QY 1007 IEDMQTEAQRBRQLLDELVALVNRKDALVRDLDAQERQABEDSHLEKTLBONKGM 1066  
 DB 921 VEDMKTEVQEKERLLESLMTTIVKRDQLVQHLNQE-TAIEDQELAKLEH--VDI 977  
 QY 1067 AKKEKCVLQ 1076  
 DB 978 SAKKDKCVLQ 987

Search completed: May 20, 2006, 17:14:41  
 Job time : 332 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2006, 17:05:37 ; Search time 199 Seconds  
(without alignments)  
2472.185 Million cell updates/sec

Title: US-09-856-723a-8  
Perfect score: 5489  
Sequence: 1 MKQYASPMPTQTDVTKFKFP.....RTLBNKGKMAKKEKCVLQ 1076

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_8.\*  
1: geneseqp1980s.\*  
2: geneseqp1990s.\*  
3: geneseqp2000s.\*  
4: geneseqp2001s.\*  
5: geneseqp2002s.\*  
6: geneseqp2003as.\*  
7: geneseqp2003bs.\*  
8: geneseqp2004s.\*  
9: geneseqp2005s.\*  
10: geneseqp2006s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5486	99.9	1076	AAU74355	AAU74355 Human cyt
2	5486	99.9	1076	ADL13160	ADL13160 Human ste
3	5442.5	99.2	1231	ADX07701	ADX07701 Cyclin-de
4	5266	95.9	1040	ADL15640	ADL15640 Human str
5	5183	94.4	1581	AAW78788	AAW78788 Human pro
6	5173	94.2	1023	AAW79773	AAW79773 Human pro
7	5173	94.2	1023	AAW79772	AAW79772 Human pro
8	5173	94.2	1023	ABG22883	ABG22883 Novel hum
9	5152	93.9	1023	AAW41108	AAW41108 Human pol
10	5152	93.9	1023	AAW41108	AAW41108 Human pol
11	5152	93.9	1023	AAW41108	AAW41108 Human pol
12	5152	93.9	1023	AAW41108	AAW41108 Human pol
13	5033	91.7	996	AAW39322	AAW39322 Human pol
14	4817	87.8	960	AAW20269	AAW20269 Novel hum
15	4485	81.7	882	ADJ70136	ADJ70136 Human hea
16	3151	57.4	669	ABG22882	ABG22882 Novel hum
17	2625	47.8	523	AAV51415	AAV51415 Human wil
18	2405	43.8	487	AAV51417	AAV51417 Human pke
19	2400	43.7	521	AAW78789	AAW78789 Human pro
20	2374	43.3	481	AAV51416	AAV51416 Human pke
21	1045	19.0	221	ABG22880	ABG22880 Novel hum
22	1020.5	18.6	957	ABW70828	ABW70828 Drosophila
23	1009	18.4	202	AAW92539	AAW92539 Human pro

24	989	18.0	479	4	ABG22881	ABG22881 Novel hum
25	945.5	17.2	1523	6	ABP55396	ABP55396 Human MOD
26	924	16.8	1452	7	ADC31335	ADC31335 Human nov
27	897	16.3	175	4	AAW17227	AAW17227 Peptide #
28	897	16.3	175	4	ABW6228	ABW6228 Peptide #
29	897	16.3	175	4	ABW31040	ABW31040 Peptide #
30	897	16.3	175	4	AAW69395	AAW69395 Human bon
31	897	16.3	175	4	AAW57004	AAW57004 Human bra
32	897	16.3	175	4	ABW51067	ABW51067 Human liv
33	897	16.3	175	4	AAW04922	AAW04922 Peptide #
34	897	16.3	175	5	ABG39012	ABG39012 Human pep
35	897	16.3	175	8	ABW60197	ABW60197 Human gen
36	510.5	9.3	996	4	ABW65591	ABW65591 Drosophila
37	507.5	9.2	1010	7	ADW48312	ADW48312 Drosophila
38	498.5	9.1	863	6	ABW98856	ABW98856 Human str
39	498.5	9.1	863	7	ADC31071	ADC31071 Human nov
40	498.5	9.1	863	7	ADW48296	ADW48296 Human MIC
41	498.5	9.1	863	8	ABW82355	ABW82355 Tumour-ag
42	493.5	9.0	863	7	ADW48308	ADW48308 Human MIC
43	473.5	8.6	398	4	AAW00864	AAW00864 Human bon
44	473.5	8.6	398	7	ADC31005	ADC31005 Human nov
45	461	8.4	91	4	AAW16170	AAW16170 Peptide #

ALIGNMENTS

RESULT 1  
AAU74355  
AAU74355 standard; protein; 1076 AA.  
AC AAU74355;  
DT 12-MAR-2002 (first entry)  
DE Human cytoskeleton-associated protein (CYSP) #26.  
KW Human; cytoskeleton-associated protein; CYSP; autoimmune disorder;  
KW cell proliferative disorder; inflammatory disorder; prion disease;  
KW vesicle trafficking disorder; gastrointestinal disorder; muscle disorder;  
KW neurological disorder; cell motility disorder; reproductive disorder;  
KW spinal cord disease; central nervous system disorder; mental disorder;  
KW gene therapy; cancer.  
OS Homo sapiens.  
PN WO200185942-A2.  
XX 15-NOV-2001.  
PF 03-MAY-2001; 2001WO-US014355.  
PR 05-MAY-2000; 2000US-0201960P.  
PR 08-MAY-2000; 2000US-0202729P.  
PR 05-JUN-2000; 2000US-0209705P.  
PR 07-JUN-2000; 2000US-0210149P.  
PR 21-JUN-2000; 2000US-0213215P.  
PA (INCY-) INCYTE GENOMICS INC.  
XX Yue H, Tang YT, Au-Young J, Lu DM, Baughn MR, Hillman JL;  
PI Azimzai Y, Lai P, Yao MG, Bandman O, Burford N, Batra S, Kearney L;  
PI Policy UJ;  
PI N-PSDB; AAS99915.  
DR WPI; 2002-062248/08.  
XX New cytoskeleton-associated proteins and polynucleotides, useful for  
PT diagnosing, preventing and treating cell proliferative, autoimmune,  
PT inflammatory, neurological, cell motility, reproductive and muscle  
PT disorders.  
XX Claim 1; Page 153-155; 194pp; English.

XX The invention relates to human cytoskeleton-associated polypeptides  
 CC (CYSKP) and their associated polynucleotide sequences. The sequences are  
 CC useful in the treatment of disorders associated with overexpression or  
 CC underexpression of CYSKP in a patient. The disorders include cell  
 CC proliferative disorders (such as cancer, actinic keratosis,  
 CC actinomyositis, cirrhosis, hepatitis and psoriasis),  
 CC autoimmune/inflammatory disorders (such as, asthma, atherosclerosis,  
 CC osteoporosis, Crohn's disease, rheumatoid arthritis, diabetes mellitus  
 CC and anaemia), vesicle trafficking disorders (such as  
 CC hypercholesterolaemia, diabetes insipidus, Grave's disease and goitre),  
 CC gastrointestinal disorders, prion diseases, neurological disorders (such  
 CC as epilepsy, stroke, cerebral neoplasms, Alzheimer's disease,  
 CC Huntington's disease, Parkinson's disease, amyotrophic lateral sclerosis  
 CC and other motor neuron disorders), cell motility disorders, reproductive  
 CC disorders (such as endometriosis and polycystic ovary syndrome), muscle  
 CC disorders (such as myocardiitis, migraine, hypertension, hypoglycaemia,  
 CC myocardial infarction, epilepsy and muscular dystrophy), spinal cord  
 CC diseases, central nervous system disorders (such as Down syndrome and  
 CC cerebral palsy) and mental disorders (such as anxiety and schizophrenia).  
 CC Sequences AAU74330-AAU74363 represent human CYSKP of the invention.

XX Sequence 1076 AA:

Query Match 99.9%; Score 5486; DB 5; Length 1076;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1075; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKQYASMPPTQTDVYKLRKPLSKRVVSALQPSISCFLEKGTDEDMOSLASVSMQ 60  
 DB 1 MKQYASMPPTQTDVYKLRKPLSKRVVSALQPSISCFLEKGTDEDMOSLASVSMQ 60  
 QY 61 ADIGNLDFEEDNDDENRNOEKAKITELINKLNFLEAKDKDLATVNSNPFDDPA 120  
 DB 61 ADIGNLDFEEDNDDENRNOEKAKITELINKLNFLEAKDKDLATVNSNPFDDPA 120  
 QY 121 AELNPFCDPSEBEITETASPRKTEDSFYNSYNPFKSVQTPQYLNPFDEBAFVTKDS 180  
 DB 121 AELNPFCDPSEBEITETASPRKTEDSFYNSYNPFKSVQTPQYLNPFDEBAFVTKDS 180  
 QY 181 PPGSTKRNKIRPVDMSKYLADSKTEBEELDESNPFYEPKSTPPNNLVNVOELETTER 240  
 DB 181 PPGSTKRNKIRPVDMSKYLADSKTEBEELDESNPFYEPKSTPPNNLVNVOELETTER 240  
 QY 241 RVRKAPAPPVLSPTKGVLENNTVSAGKDLSTSPSPSPSPSPSPSPSPSPSPSPSPSP 300  
 DB 241 RVRKAPAPPVLSPTKGVLENNTVSAGKDLSTSPSPSPSPSPSPSPSPSPSPSPSPSP 300  
 QY 301 VTKNVRGVKLTNFTTSWRNGLSFCAILHFRPDLIDYKSLNPPDIKENNKAYDGFASIG 360  
 DB 301 VTKNVRGVKLTNFTTSWRNGLSFCAILHFRPDLIDYKSLNPPDIKENNKAYDGFASIG 360  
 QY 361 ISRLLEPSDMVLAIIPDKLTVMYLYOIRAHFSGQELNVVQIEBNSKSTYKGVNEDT 420  
 DB 361 ISRLLEPSDMVLAIIPDKLTVMYLYOIRAHFSGQELNVVQIEBNSKSTYKGVNEDT 420  
 QY 421 NSGVDOEKFYAEISDLKREBELQPIGAVDFLSODSVFVNSGVGESSEHQTDPDHL 480  
 DB 421 NSGVDOEKFYAEISDLKREBELQPIGAVDFLSODSVFVNSGVGESSEHQTDPDHL 480  
 QY 481 SPSTASPYCARTKSDTEPQKSSQSGRTSGSDPGICSNSTDSYQAVLLGKRLLKMETL 540  
 DB 481 SPSTASPYCARTKSDTEPQKSSQSGRTSGSDPGICSNSTDSYQAVLLGKRLLKMETL 540  
 QY 541 ELSDLYVSDKKKQMSPPFICEETDEQKLTIDIGSNIEKELNSRLSECRSDPESPIKK 600  
 DB 541 ELSDLYVSDKKKQMSPPFICEETDEQKLTIDIGSNIEKELNSRLSECRSDPESPIKK 600  
 QY 601 TSLSPSTSKGYSTRDLAKKKGASLRQTESDDPADRTLLNHADHSKIVQRLLSROE 660  
 DB 601 TSLSPSTSKGYSTRDLAKKKGASLRQTESDDPADRTLLNHADHSKIVQRLLSROE 660  
 QY 661 ELKERARVLLQEARDAALAKGNKNTNTATPFCNRQLSDQODERRRQLERARQLIAE 720

DB 661 ELKERARVLLQEARDAALAKGNKNTNTATPFCNRQLSDQODERRRQLERARQLIAE 720  
 QY 721 ARSGVMSKSLPSYGEAAAEKLSKASGDENDNIEIDTNEELPEGVVGGBDELTVLEN 780  
 DB 721 ARSGVMSKSLPSYGEAAAEKLSKASGDENDNIEIDTNEELPEGVVGGBDELTVLEN 780  
 QY 781 DLDPPEONSKLVLDKLLKLEVPQVANSPSAAQKAVTSSSEODMSGTEDLRTERLQK 840  
 DB 781 DLDPPEONSKLVLDKLLKLEVPQVANSPSAAQKAVTSSSEODMSGTEDLRTERLQK 840  
 QY 841 TTERFRNPVVSFSDSTYRKTQLOFSQYIBNRPKMRQSIQEDTKKNEEKAITETOR 900  
 DB 841 TTERFRNPVVSFSDSTYRKTQLOFSQYIBNRPKMRQSIQEDTKKNEEKAITETOR 900  
 QY 901 KPSEDEVLNKGFKUTSQYVGEALALENEKQIDTRALVBRKRLYIMDTGRNTEESBA 960  
 DB 901 KPSEDEVLNKGFKUTSQYVGEALALENEKQIDTRALVBRKRLYIMDTGRNTEESBA 960  
 QY 961 MOEWFMVLNKKNALIRMMQSLLEKEHDLERYEELNRELRAMLAIEDMQTEAOKRE 1020  
 DB 961 MOEWFMVLNKKNALIRMMQSLLEKEHDLERYEELNRELRAMLAIEDMQTEAOKRE 1020  
 QY 1021 QLLLDLDELVALVNRKDALVRDLDAQEKQAEDEDEHLETLTBOBKGAKEKCVLQ 1076  
 DB 1021 QLLLDLDELVALVNRKDALVRDLDAQEKQAEDEDEHLETLTBOBKGAKEKCVLQ 1076

RESULT 2  
 ADLJ3160  
 ID ADLJ3160 standard; protein; 1076 AA.  
 AC ADLJ3160;  
 XT 06-MAY-2004 (first entry)  
 XX Human steroid-induced C3A liver cell protein #134.  
 XX Hepatotropic; Gene therapy; Wilson disease; liver disorder;  
 KW steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.  
 OS Homo sapiens.  
 PN US6673549-B1.  
 PD 06-JAN-2004.  
 XX 12-OCT-2001; 2001US-00976594.  
 PF 12-OCT-2000; 2000US-0240409P.  
 PR 12-OCT-2000; 2000US-0240409P.  
 XX (INCYT-) INCYTE CORP.  
 PA Furness LM, Buchbinder JL;  
 PI WPI; 2004-068610/07.  
 DR Combination useful for preparing a composition for treating liver  
 XX disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,  
 PT comprises cDNAs that are differentially expressed in response to steroid  
 PT treatment.  
 XX Disclosure; SEQ ID NO 889; 141pp; English.  
 XX The invention relates to a combination comprising cDNAs that are  
 CC differentially expressed in response to steroid treatment. Also included  
 CC are the following: a high throughput method for using a cDNA to detect  
 CC differential expression of nucleic acids in a sample; and a high  
 CC throughput method of screening molecules or compounds to identify a  
 CC ligand that specifically binds a cDNA. The sample is from a subject with  
 CC Wilson disease and comparison of a standard defines a stage of that  
 CC disease. The high throughput method of screening molecules or compounds  
 CC to identify a ligand that specifically binds a cDNA comprises: combining

CC the combination with molecules or compounds under conditions to allow  
CC specific binding; and detecting specific binding between each cDNA and at  
CC least one molecule or compound. The molecules or compounds are regulatory  
CC proteins. The combination is useful for preparing a composition for  
CC treating liver disorders associated with steroid therapy, e.g., cirrhosis  
CC or hepatitis. The present sequence represents a human protein which is  
CC differentially expressed in steroid-induced C3A liver cells. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format directly from USPTO  
CC at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

XX Sequence 1076 AA;

Query Match 99.9%; Score 5486; DB 8; Length 1076;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1075; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKQYASPMPTQTDVTKLKEPLSKRVSAALQFSLSCIFLRBGAATDEDMQSLASLVSMQ 60  
DB 1 MKQYASPMPTQTDVTKLKEPLSKRVSAALQFSLSCIFLRBGAATDEDMQSLASLVSMQ 60  
QY 61 ADIGNLDFEEDNEDDENRVQAEKAKITELINKLFLDEABKOLATVSNPPDDDA 120  
DB 61 ADIGNLDFEEDNEDDENRVQAEKAKITELINKLFLDEABKOLATVSNPPDDDA 120  
QY 121 AELNPFQDSEBEPITETSPKTEDSFYNNSYNPFKEVQTPQVLANPDEPAFTIKDS 180  
DB 121 AELNPFQDSEBEPITETSPKTEDSFYNNSYNPFKEVQTPQVLANPDEPAFTIKDS 180  
QY 181 PPGSTKRKNIRPVMSKYLVAADSKTEBEBELDESNPFEYKSTPPNNLVNPFQLETER 240  
DB 181 PPGSTKRKNIRPVMSKYLVAADSKTEBEBELDESNPFEYKSTPPNNLVNPFQLETER 240  
QY 241 RYKRAAPAPVLSPTKGLVLENTVYSAKOLSTSPSPSPSPSPSPSPSPSPSPSPSPSP 300  
DB 241 RYKRAAPAPVLSPTKGLVLENTVYSAKOLSTSPSPSPSPSPSPSPSPSPSPSPSPSP 300  
QY 301 VTQKRYGKVTNFTSMWNGLSFCALHFRPDLIDYKSLNPQDIKENNKAYDGFASIG 360  
DB 301 VTQKRYGKVTNFTSMWNGLSFCALHFRPDLIDYKSLNPQDIKENNKAYDGFASIG 360  
QY 361 ISRLLEPSDMVLAIPDKLVTWYLYQIRAHFSGGELNVQIEENSSTKYKVGNEYEDT 420  
DB 361 ISRLLEPSDMVLAIPDKLVTWYLYQIRAHFSGGELNVQIEENSSTKYKVGNEYEDT 420  
QY 421 NSSVDQEFYALSDLKREPELQPIGAVDFLSGDDSVFVNDSCVGSSESHQTPDDHL 480  
DB 421 NSSVDQEFYALSDLKREPELQPIGAVDFLSGDDSVFVNDSCVGSSESHQTPDDHL 480  
QY 481 SPSTASPCRRTPKSTEPKQSSGRTSGSDPGICNTSTQVLAQVLLGKKRLKATL 540  
DB 481 SPSTASPCRRTPKSTEPKQSSGRTSGSDPGICNTSTQVLAQVLLGKKRLKATL 540  
QY 541 ELSDLYVADKKKMSPPICEETDEQKQTLIDIGSNLEKEKLENSRLSECRSDPSPIKX 600  
DB 541 ELSDLYVADKKKMSPPICEETDEQKQTLIDIGSNLEKEKLENSRLSECRSDPSPIKX 600  
QY 601 TSLSTFSLGYSYSDLDLAKKKASLQTESDPDADRITINHADSHSKIVQHRLLSQE 660  
DB 601 TSLSTFSLGYSYSDLDLAKKKASLQTESDPDADRITINHADSHSKIVQHRLLSQE 660  
QY 661 ELKERARVLLEQARDAALAKGNKNTNTATPPFCRQSLSDQDERRRQALERRAQILAE 720  
DB 661 ELKERARVLLEQARDAALAKGNKNTNTATPPFCRQSLSDQDERRRQALERRAQILAE 720  
QY 721 ARSGVMSKELPSYGEAAAEKLEKSKAGDENNDIETNEIIPGFVVGDELTNLEN 780  
DB 721 ARSGVMSKELPSYGEAAAEKLEKSKAGDENNDIETNEIIPGFVVGDELTNLEN 780  
QY 781 DLDTPEONSKIVDLKLLKLEVOQVANSPPSSAAQKAVTESSEBODMKSCTEDLTERLOK 840  
DB 781 DLDTPEONSKIVDLKLLKLEVOQVANSPPSSAAQKAVTESSEBODMKSCTEDLTERLOK 840

QY 841 TTERFRNPVVSQDSTVKTQLOSFQYIENRPEMKRORSIOEDYKNGEKAATETOR 900  
DB 841 TTERFRNPVVSQDSTVKTQLOSFQYIENRPEMKRORSIOEDYKNGEKAATETOR 900  
QY 901 KPSEDEVANKGFKTQSYVVGELALLENQKOIDTRALVKKRLYLMDTGNTSEBEM 960  
DB 901 KPSEDEVANKGFKTQSYVVGELALLENQKOIDTRALVKKRLYLMDTGNTSEBEM 960  
QY 961 MOEWMFLVNNKQALIRRNQSLLEKEDLRRRYELNRELRAMLAIDMOKTEAKORE 1020  
DB 961 MOEWMFLVNNKQALIRRNQSLLEKEDLRRRYELNRELRAMLAIDMOKTEAKORE 1020  
QY 1021 QLLDELVALVNNKDALVRDLDQAEKQAEDEHLERTLBOCKGMAKKEKCVIQ 1076  
DB 1021 QLLDELVALVNNKDALVRDLDQAEKQAEDEHLERTLBOCKGMAKKEKCVIQ 1076

RESULT 3  
ADK07701  
ID ADK07701 standard; protein; 1231 AA.  
XX  
AC ADK07701;  
XX  
DT 21-APR-2005 (first entry)  
XX  
DE Cyclin-dependent kinase modulation biomarker SEQ ID NO 2266.  
XX  
KM cytostatic; cyclin-dependent kinase; cdk; biomarker.  
XX  
OS Homo sapiens.  
XX  
PN WO2005012875-A2.  
XX  
PD 10-FEB-2005.  
XX  
PF 29-JUL-2004; 2004WO-US024424.  
XX  
PR 29-JUL-2003; 2003US-0430890P.  
XX  
PA (BRIM) BRISTOL-MYERS SQUIBB CO.  
XX  
PI Li M, Rudnow BA, Webster KR, Jackson DG, Wong TW,  
XX  
DR WPI; 2005-163068/17.  
XX  
DR N-P8DB; ADK07700.  
XX  
PT Biomarkers useful for predicting or determining the response of a mammal  
XX  
PT to a cancer treatment comprising administration of a modulator of cyclin-  
XX  
PT dependent kinase activity.  
XX  
PS Claim 5; SEQ ID NO 2266; 141bp; English.  
XX  
CC This invention describes a novel method of predicting or determining  
CC whether a mammal will respond or is responding to an anti-cancer agent  
CC that modulates cyclin-dependent kinase (cdk) activity. The method  
CC comprises measuring the level of one or more biomarkers selected from  
CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID  
CC NO.1246 (Genbank E87 W28729) is especially preferred). The method of the  
CC invention is utilized in a kit for determining or predicting whether  
CC patient would be susceptible or resistant to treatment by an agent  
CC modulating cdk activity. The invention also describes a method for  
CC utilizing individualized genetic profiles for treating diseases and  
CC disorders based on patient's response and molecular level, specialized  
CC microarrays comprising the biomarkers described, antibodies directed  
CC against the biomarkers and a cell culture model to identify biomarkers.  
CC The cdk modulator is preferably N-5-[(5-(1,1-dimethylethyl)-2-  
CC oxazolyl)methylthio]-2-thiazolyl-4-piperidine carboxamide, 0.5-1-  
CC tartaric acid salt. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences). This  
CC sequence represents a biomarker used in the method of the invention.  
XX  
XX Sequence 1231 AA;

Query Match 99.2%; Score 5442.5; DB 9; Length 1231;  
 Best Local Similarity 96.6%; Pred. No. 0;  
 Matches 1073; Conservative 1; Mismatches 2; Indels 35; Gaps 1;

QY 1 MKQYASMPPTOTDYKFKPKLSKVVSAAALQPSISCIPLREGKATDEBMOSLASVSKQ 60  
 DB 121 MKQYASMPPTOTDYKFKPKLSKVVSAAALQPSISCIPLREGKATDEBMOSLASVSKQ 180  
 QY 61 ADIGNLDPEDNEDDENRYNOBEKAKIT----- 91  
 DB 181 ADIGNLDPEDNEDDENRYNOBEKAKITGIVNQLNALSLLDEDODDCIKQANMSAK 240  
 QY 92 -----ELINKANLEDAEKOLATVNSNPFDDPDAAELNPGDDPDESEPIETASPKTB 145  
 DB 241 SASSEELINKANLEDAEKOLATVNSNPFDDPDAAELNPGDDPDESEPIETASPKTB 300  
 QY 146 DSYFNNSYNPPEKEVQTPQYINPDEPEAFVTIKDSPPOSTRKNIIPVMSKYLYADSSK 205  
 DB 301 DSYFNNSYNPPEKEVQTPQYINPDEPEAFVTIKDSPPOSTRKNIIPVMSKYLYADSSK 360  
 QY 206 TEEBELESNPFYKSTPPNNLVNPQLETERAVYKRAPAPVLSPTKGVANENTVS 265  
 DB 361 TEEBELESNPFYKSTPPNNLVNPQLETERAVYKRAPAPVLSPTKGVANENTVS 420  
 QY 266 AGKDLSTSPKPSPIPSVPLGRKPNASOSILVCKEVTNKYGVKITFTTSMRAGLSFCA 325  
 DB 421 AGKDLSTSPKPSPIPSVPLGRKPNASOSILVCKEVTNKYGVKITFTTSMRAGLSFCA 480  
 QY 326 ILHHFRPDLIDYKSLNPODIKENNKAAYDGFASIGISRLLEPSDMVLAIIPDKLTVMTYL 385  
 DB 481 ILHHFRPDLIDYKSLNPODIKENNKAAYDGFASIGISRLLEPSDMVLAIIPDKLTVMTYL 540  
 QY 386 YQIRAHSSGOLNVQIENSSKSTYKVNGETPTNSSVDDEKYAELSDIKREPELOP 445  
 DB 541 YQIRAHSSGOLNVQIENSSKSTYKVNGETPTNSSVDDEKYAELSDIKREPELOP 600  
 QY 446 ISGAVDELSDODSVFVNDGVSSESSEHQTPDDHLSPTASPYCRRTKSDTEPOKSOQS 505  
 DB 601 ISGAVDELSDODSVFVNDGVSSESSEHQTPDDHLSPTASPYCRRTKSDTEPOKSOQS 660  
 QY 506 GRTSGSDPGICSNSTDTQAQVLLGKRLKAEFTLELSDLVYSDKKKOMPPICEETDE 565  
 DB 661 GRTSGSDPGICSNSTDTQAQVLLGKRLKAEFTLELSDLVYSDKKKOMPPICEETDE 720  
 QY 566 OKQOTLDIGSNLEKKEKLENSSLERCRSDPESPIKKTSLSPYSKLGYSYRDLIAKKHA 625  
 DB 721 OKQOTLDIGSNLEKKEKLENSSLERCRSDPESPIKKTSLSPYSKLGYSYRDLIAKKHA 780  
 QY 626 SLRQTESDPADRTTLNHAHSSKIYHRLLSROBELKERARVLLBOARRDALAKAGKH 685  
 DB 781 SLRQTESDPADRTTLNHAHSSKIYHRLLSROBELKERARVLLBOARRDALAKAGKH 840  
 QY 686 NNTATPFCNRQJSDQDEERRRQJREARQJLAARSGVMSLEPSYGEMAAEKLKERS 745  
 DB 841 NNTATPFCNRQJSDQDEERRRQJREARQJLAARSGVMSLEPSYGEMAAEKLKERS 900  
 QY 746 KASGDENDNIEIDNNEIPEGFVVGSDDELNLNENDLDPONSKLYDLKUKLLSVQPO 805  
 DB 901 KASGDENDNIEIDNNEIPEGFVVGSDDELNLNENDLDPONSKLYDLKUKLLSVQPO 960  
 QY 806 VANSPPSAQAQVATESEODMKSGTEDLARTERLOKTERFPNPNVVFESKDSVTRKTOQSF 865  
 DB 961 VANSPPSAQAQVATESEODMKSGTEDLARTERLOKTERFPNPNVVFESKDSVTRKTOQSF 1020  
 QY 866 SQTENRPEMKRQSIQEDTKKGBEKAITETQRKSEDEVINKGFQDTSQYVVGELAA 925  
 DB 1021 SQTENRPEMKRQSIQEDTKKGBEKAITETQRKSEDEVINKGFQDTSQYVVGELAA 1080  
 QY 926 LBNBOKQIDTRALVERKRLYIMDTGNTBEEBAMQEWPLVKKKALLRRNQJLSLE 985  
 DB 1081 LBNBOKQIDTRALVERKRLYIMDTGNTBEEBAMQEWPLVKKKALLRRNQJLSLE 1140

QY 986 KEHDLERREYELNRELRAMLAIEDMQTEAQKREBOLLDELVALVVKRDALVEDLDAQE 1045  
 DB 1141 KEHDLERREYELNRELRAMLAIEDMQTEAQKREBOLLDELVALVVKRDALVEDLDAQE 1200  
 QY 1046 KOAREDEBHLERTLEONKGMKAKKEKCVIQ 1076  
 DB 1201 KOAREDEBHLERTLEONKGMKAKKEKCVIQ 1231

RESULT 4  
 ADE15640  
 ID ADE15640 standard; protein; 1040 AA.  
 XX  
 AC ADE15640;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Human structural and cytoskeleton-associated protein (SCAP) #19.  
 XX  
 KW human; structural and cytoskeleton-associated protein; SCAP;  
 KW arteriosclerosis; atherosclerosis; cirrhosis; hepatitis; myelofibrosis;  
 KW psoriasis; cancer; pneumonia; chronic bronchitis; yellow fever;  
 KW influenza; measles; mumps; HIV; human T lymphotropic virus; rabies;  
 KW gastroenteritis; encephalitis; rubella; epilepsy;  
 KW ischaemic cerebrovascular disease; stroke; cerebral neoplasm;  
 KW Alzheimer's disease; Pick's disease; Huntington's disease; dementia;  
 KW Parkinson's disease; amyotrophic lateral sclerosis; atrophy;  
 KW hereditary ataxia; multiple sclerosis; meningitis; brain abscess;  
 KW prion disease; Creutzfeldt-Jakob disease; Insomnia; neurofibromatosis;  
 KW cerebral palsy; myasthenia gravis; anxiety.  
 XX  
 OS Homo sapiens.  
 XX  
 XX W02003062391-A2.  
 XX  
 PD 31-JUL-2003.  
 XX  
 PF 16-JAN-2003; 2003MO-US001772.  
 XX  
 PR 18-JAN-2002; 2002US-0350702P.  
 PR 25-JAN-2002; 2002US-0351715P.  
 PR 15-FEB-2002; 2002US-0357402P.  
 PR 10-MAY-2002; 2002US-0379880P.  
 PR 17-MAY-2002; 2002US-0381599P.  
 PR 07-JUN-2002; 2002US-0387270P.  
 PR 19-JUL-2002; 2002US-0397125P.  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 XX  
 PI Yue H, Griffin JA, Richardson TW, Tang YT, Thangaveju K;  
 PI Foreyche J, Becha SD, Chavla NK, Hafajia AJ, Swarnakar A;  
 PI Marguis JP, Gotved AE, Baughn MR, Lu DM, Arizizu CS, Kable AE;  
 PI Lee ST, Ramkumar J, Jjiang X, Jackson AA, Khare R, Elliott VS;  
 PI Bulloch SA, Xu Y, Lee S, Lehr-Mason PM;  
 XX  
 DR WPI; 2003-671468/63.  
 DR N-PSDB; ADE15671.  
 XX  
 PT New isolated polypeptides useful for treating e.g. cell proliferative  
 PT disorders, viral infections and neurological disorders.  
 XX  
 PS Claim 1; SEQ ID NO 19; 357pp; English.  
 XX  
 CC The invention comprises the amino acid and coding sequences of human  
 CC structural and cytoskeleton-associated proteins (SCAP). The SCAP DNA and  
 CC protein sequences of the invention are useful for the diagnosis and  
 CC treatment of: arteriosclerosis, atherosclerosis, cirrhosis, hepatitis,  
 CC myelofibrosis, psoriasis, primary cancer, pneumonia, chronic bronchitis,  
 CC yellow fever, influenza, measles, mumps, HIV, human T lymphotropic virus,  
 CC rabies, gastroenteritis, encephalitis, rubella, epilepsy, ischaemic  
 CC cerebrovascular disease, stroke, cerebral neoplasm, Alzheimer's disease,  
 CC Pick's disease, Huntington's disease, dementia, Parkinson's disease,  
 CC amyotrophic lateral sclerosis, atrophy, hereditary ataxias, multiple

CC scierosis, meningitis, brain abscess, Creutzfeldt-Jakob  
CC disease, insomnia, neurofibromatosis, cerebral palsy, myasthenia gravis  
CC anxiety. The present amino acid sequence represents a human SCAP of the  
CC invention.

**SQ** Sequence 1040 AA;

Query Match	95.9%	Score 5266;	DB 7;	Length 1040;
Best Local Similarity	96.6%;	Pred. No. 2.8e-316;		
Matches 1039;	Conservative	1;	Mismatches 0;	Indels 36; Gaps 1

Qy	1	MQYASMPQJVDVJLKFPLSKVYSALQSLSCIFLREKADDEMOSIALVSMKQ	60
Db	1	MQYASMPQJVDVJLKFPLSKVYSALQSLSCIFLREKADDEMOSIALVSMKQ	60
Qy	61	ADJGNTLDDPEEDNEDDERVNOEKAATITELINCLNPLDEAKOLATVNSNPPDDPA	1200
Db	61	ADJGNTLDDPEEDNEDDERVNOEKAATITELINCLNPLDEAKOLATVNSNPPDDPA	1200
Qy	121	ABINPFGDDPDSSEPIETETASPKRTBDSFNNSYNPFKVCQPOYANLPDEBEAVTJKDS	1800
Db	121	ABINPFGDDPDSSEPIETETASPKRTBDSFNNSYNPFKVCQPOYANLPDEBEAVTJKDS	1800
Qy	181	PPGSTRKKNIRPVDMSKYLADSSKTEBEHLDESNPFEYKSTPEPPNVLVNPVOELETER	2400
Db	181	PPGSTRKKNIRPVDMSKYLADSSKTEBEHLDESNPFEYKSTPEPPNVLVNPVOELETER	2400
Qy	241	RVRKKA.PAPPVLSPTKTVLNTENTVSAQDLSISPYCPSP1PSPVJGRKNASQSLVWCX	3000
Db	241	RVRKKA.PAPPVLSPTKTVLNTENTVSAQDLSISPYCPSP1PSPVJGRKNASQSLVWCX	3000
Qy	301	VTGNVGVKLTNTTWSRNGLSFCAL1LHFRPDLIDYSLNMPDQIKENNKAYGSPASIG	3600
Db	301	VTGNVGVKLTNTTWSRNGLSFCAL1LHFRPDLIDYSLNMPDQIKENNKAYGSPASIG	3600
Qy	361	ISRLLEPDMVLLAIPDKLTWNTYLYQIRAHFSGQELNVQIENSSKSTYKVGNYZTD	4200
Db	361	ISRLLEPDMVLLAIPDKLTWNTYLYQIRAHFSGQELNVQIENSSKSTYKVGNYZTD	4200
Qy	421	NSSVDDPKFYAELSDIKREPEIQOPISGAVPLSGDDSVFVNDSGVGBESBHQTPDHL	4800
Db	421	NSSVDDPKFYAELSDIKREPEIQOPISGAVPLSGDDSVFVNDSGVGBESBHQTPDHL	4800
Qy	481	SPSTASPYCRRTKSDTEPOKSOOSSGRTSGSDPDCSNTSTOQAQVYLGGKRL1KAEVL	5400
Db	481	SPSTASPYCRRTKSDTEPOKSOOSSGRTSGSDPDCSNTSTOQAQVYLGGKRL1KAEVL	5400
Qy	541	ELSDLYVSDKKQDMSPP1CBEETDQKLTJLDSNLEKEKLENSRSLACRSDPESP1KK	6000
Db	541	ELSDLYVSDKKQDMSPP1CBEETDQKLTJLDSNLEKEKLENSRSLACRSDPESP1KK	6000
Qy	601	TLSLSPFSK1GY5SRJDLAKKGA5LRQTESDPADRTT1NHADHSG1YQHR1LSROB	6600
Db	601	TLSLSPFSK1GY5SRJDLAKKGA5LRQTESDPADRTT1NHADHSG1YQHR1LSROB	6600
Qy	661	ELKERARV1LEQARRPAALKAQNKXNTTAT1PFCRQJLSDQDEERRRQJLERARQJLAE	7200
Db	661	ELKERARV1LEQARRPAALKAQNKXNTTAT1PFCRQJLSDQDEERRRQJLERARQJLAE	7200
Qy	721	ARSGVMSBLPSYGEMAAEKLKERSKASGENDN1EIDTNEB1PBGFVGGDELTYL1EN	7800
Db	721	ARSGVMSBLPSYGEMAAEKLKERSKASG-----	7490
Qy	781	DLDTPEONSKVDL1KJLLEVOPOVANSPSAAQATVSSBQDMSGTEDLATEBLQK	8400
Db	750	-----BQNSKLVDLK1KJLLEVOPOVANSPSAAQATVSSBQDMSGTEDLATEBLQK	8040
Qy	841	TTFRFPN1PVVFSKDSIVRTQLOSFSQY1ENPPEKRRORS1QEDTJKGNBEKALITEQR	9000
Db	805	TTFRFPN1PVVFSKDSIVRTQLOSFSQY1ENPPEKRRORS1QEDTJKGNBEKALITEQR	8640
Qy	901	KQGBDEV1LNGFRDTSQYVGBELAL1ENBQO1DTRALV1BKRL1Y1LMDTGRNTEBEAM	9600

Db	865	KPSEBVLNKGFKDTSQYVWESIALLENBQKQIDTPRALVETRILRYMDTGNTSEBEM	924
Oy	961	MOEMFMLVNKKGNALIRBANO,SLIEKENDDLEERYELINRELPAMALIEDMORTBAOKRE	102
Db	925	MOEMFMLVNKKGNALIRBNOQ,SLIEKENDDLEERYELINRELPAMALIEDMORTBAOKRE	984
Oy	1021	QULDELVALVNKRDALVYDIDAOEKABEBSBHERTEJONKGNKAKKEBKVLO	1076
Db	985	QULDELVALVNKRDALVYDIDAOEKABEBSBHERTEJONKGNKAKKEBKVLO	1040

RESULT 5  
AAM78788

ID AAM78788 standard; protein; 1581 AA.

AC AAM78788;

DT 06-NOV-2001 (first entry)

DE Human protein SEQ ID NO 1450.

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW tissue growth factor; immunomodulatory; cancer; leukaemia; narrow-spectrum antibody; inflammation

**KW** nervous system disorder; arthritis; inflammation.

05 **Homo sapiens.**

**Homo sapiens**

PN WO200157190-A2

PD 09-AUG-2001

PF 05-FEB-2001; 2001WO-US004098.

PR 03-FEB-2000; 2000US-00496914.

PR 20-JUN-2000; 2000US-00598075.

PR 01-SEP-2000; 2000US-00654936.

PR 20-OCT-2000; 2000US-00693325.

2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

PI Ma Y, Zhao QA, Wang D, Wang

XX

DR N-PSDB; AAK51921.

PT Nucleic acids encoding polypeptides with cytokine-like activities, use

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----

XX

CC encoded polypeptides

CC production of oth

peptide therapy. The polypeptide

CC activity, tissue growth factor

CC treatment of cancer, leukaemia, nervous system

CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAM78923-AAH80302) that exhibit activity elating  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activity/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK5581), 2111  
CC (AAK5582) and 3666 (AAM80020) are omitted as the relevant pages from the  
CC sequence listing were missing at the time of publication

**SQ Sequence 1581 AA;**

50

Query Match 94.4%; Score 5183; DB 4; Length 1581;  
 Best Local Similarity 67.6%; Pred. No. 6.7e-311;  
 Matches 1068; Conservative 1; Mismatches 0; Indels 512; Gaps 1;

```

QY      8 MPQTQVYKLRKPLSKVYVSAALQPSLSCIFLRGSKATDEDMQSLASIVSKQADIGNLD 67
DB      1 MPQTQVYKLRKPLSKVYVSAALQPSLSCIFLRGSKATDEDMQSLASIVSKQADIGNLD 60

QY      68 DFEEDNDDDBENRNOEBKAKITELINKLNFLEBAEKDLATVNSNPDDDAELNPF 127
DB      61 DFEEDNDDDBENRNOEBKAKITELINKLNFLEBAEKDLATVNSNPDDDAELNPF 120

QY      128 DPSEBEPITETASPRKTEDSPYNNNSYNPFKEVQTPQVLPNPFDEBAFVYTIKOSPQSTKR 187
DB      121 DPSEBEPITETASPRKTEDSPYNNNSYNPFKEVQTPQVLPNPFDEBAFVYTIKOSPQSTKR 180

QY      188 KNTRPVMSKYLVDSSKTEBEELEDSNPVEYKSTPPNNLVNPVOELETERVVKAKAP 247
DB      181 KNTRPVMSKYLVDSSKTEBEELEDSNPVEYKSTPPNNLVNPVOELETERVVKAKAP 240

QY      248 APPVLSPRTGVLANENTVSAGKDLSTSPKPSPIPSPVLRGKPNASQSLVWCKEYTKQYRG 307
DB      241 APPVLSPRTGVLANENTVSAGKDLSTSPKPSPIPSPVLRGKPNASQSLVWCKEYTKQYRG 300

QY      308 VKITNFTTSMRNGLSFCALIHFRPDLIDYKSLNPODIKENNKAAYDGFASIGISRLLEP 367
DB      301 VKITNFTTSMRNGLSFCALIHFRPDLIDYKSLNPODIKENNKAAYDGFASIGISRLLEP 360

QY      368 SDWVLLAIPDKLTVMYLYOIRAHFSGOELNVQIENSSKSTYKVGNYETDTNNSVDQ 427
DB      361 SDWVLLAIPDKLTVMYLYOIRAHFSGOELNVQIENSSKSTYKVGNYETDTNNSVDQ 420

QY      428 KFYAELSDLRBPELQOPIGAVDFLSQDPSVFVNDSGVGESEHQTPODHLSPSTASP 487
DB      421 KFYAELSDLRBPELQOPIGAVDFLSQDPSVFVNDSGVGESEHQTPODHLSPSTASP 480

QY      488 YCRRTKSDTEPOKSOQSSGRTSGSDDPGICSNITDSTQAOVLGKKRLKKAETLELSDLYV 547
DB      481 YCRRTKSDTEPOKSOQSSGRTSGSDDPGICSNITDSTQAOVLGKKRLKKAETLELSDLYV 540

QY      548 SDDKKDMSPPPICETEBEQKQTLIDIGSNLEKEKLENSRSIECSDPESPICKTSLSEPTS 607
DB      541 SDDKKDMSPPPICETEBEQKQTLIDIGSNLEKEKLENSRSIECSDPESPICKTSLSEPTS 600

QY      608 KLGYSYSRDLDLAKKQKASLRQTESDPDADRTTLNHADHSSKIVQHRLLSROBELKERAR 667
DB      601 KLGYSYSRDLDLAKKQKASLRQTESDPDADRTTLNHADHSSKIVQHRLLSROBELKERAR 660

QY      668 VLEQARDAALAKGNKGNNTNTATPCNRQLSDQODERRRQLERARQLAERASGVKM 727
DB      661 VLEQARDAALAKGNKGNNTNTATPCNRQLSDQODERRRQLERARQLAERASGVKM 720

QY      728 SELPSYGEAAEKLEKSKASGDENDNIEIDTNEEIPBEGVYGGDELTNLENDLQTEQ 787
DB      721 SELPSYGEAAEKLEKSKASGDENDNIEIDTNEEIPBEGVYGGDELTNLENDLQTEQ 780

QY      788 NSKLVDLKLKKLLEVOQVANSPPSSAOKAVTESSEDMKSGTEDLATERLOKTERFRN 847
DB      781 NSKLVDLKLKKLLEVOQVANSPPSSAOKAVTESSEDMKSGTEDLATERLOKTERFRN 840

QY      848 PVVFSKDSIVRKTQLOSFQYIENRPMKORSIOEDTKKGNBEKKAITETORKPSDEV 907
DB      841 PVVFSKDSIVRKTQLOSFQYIENRPMKORSIOEDTKKGNBEKKAITETORKPSDEV 900

QY      908 LNKGFQDTSQYVVGELALENEQKQIDTRALVEKRLRYLMDTGRNTEEBAMQEWFM 967
DB      901 LNKGFQDTSQYVVGELALENEQKQIDTRALVEKRLRYLMDTGRNTEEBAMQEWFM 960

QY      968 VNNKNLIRMNQSLLEKEHDLERRRYELNR----- 999
DB      961 VNNKNLIRMNQSLLEKEHDLERRRYELNR----- 999

QY      1000 ----- 999
  
```

```

DB      1021 IRAHFGOELNVQIENSSKSTYKVGNYETDTNNSVDQEKFYAELSDLRBPELQOPI 1060
QY      1000 ----- 999
DB      1081 GAVDFLSQDPSVFVNDSGVGESEHQTPODHLSPSTASPCRTKSDTEPOKSOSSGR 1140
QY      1000 ----- 999
DB      1141 TSGSDDPGICSNITDSTQAOVLGKKRLKKAETLELSDLYVSDKKDMSPPPICETEBEQ 1200
QY      1000 ----- 999
DB      1201 LQTLIGSNLEKEKLENSRSIECSDPESPICKTSLSPSKLGYSYSRDLDLAKKQKASL 1260
QY      1000 ----- 999
DB      1261 RQTESDPDADRTTLNHADHSSKIVQHRLLSROBELKERARVLEQARDAALAKGNKNT 1320
QY      1000 ----- 999
DB      1321 NTATPCNRQLSDQODERRRQLERARQLAERASGVKMSLPSYGEAAEKLEKSKA 1380
QY      1000 ----- 999
DB      1381 SGEQNSKLVLDLKLKKLLEVOQVANSPPSSAOKAVTESSEDMKSGTEDLATERLOKTE 1440
QY      1000 ----- 999
DB      1441 RFRNPVVSFSDSTVRKTQLOSFQYIENRPMKORSIOEDTKKGNBEKKAITETORPS 1500
QY      1000 -----ELRAMLAIEDMOKTEBAQKRREQLLDELVALVVKRDALVRDLDQAKOABEEDHL 1055
DB      1501 EDEVELRAMLAIEDMOKTEBAQKRREQLLDELVALVVKRDALVRDLDQAKOABEEDHL 1560
QY      1056 ERTLQNKGMKAKKEKCVLQ 1076
DB      1561 ERTLQNKGMKAKKEKCVLQ 1581

RESULT 6
AAM79773
ID AAM79773 standard; protein; 1023 AA.
XX
AC AAM79773;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 3419.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN K0200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001MO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSB-) HYSEQ INC.
  
```





PS Claim 20, Page 328; 6221pp; English.  
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAW78323-AAW80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
CC (AAK52582) and 3666 (AAW80020) are omitted as the relevant pages from the  
CC sequence listing were missing at the time of publication  
XX  
SQ Sequence 1023 AA;  
Query Match 94.2%; Score 5173; DB 4; Length 1023;  
Best Local Similarity 99.7%; Pred. No. 1.6e-310;  
Matches 1011; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 63 IGNUIDFEDNEEDDENRVNOEKAATBELINKLNFIDEAKDLATNSNPDPDAE 122  
DB :  
DB 10 LANIDDEBEDNEEDDENRVNOEKAATBELINKLNFIDEAKDLATNSNPDPDAE 69  
QY 123 LNPFGDDPDEEPIETETSPKTEBSFYNNSYNPFGEVQTPQYLNPDPDEEAFVTIKDSP 182  
DB LNPFGDDPDEEPIETETSPKTEBSFYNNSYNPFGEVQTPQYLNPDPDEEAFVTIKDSP 129  
QY 183 QSTGRKRIKPYDMKSYLYADSSKTEEBELDESNPFYEKSTPPNNLVNPVQELETERRV 242  
DB 130 QSTGRKRIKPYDMKSYLYADSSKTEEBELDESNPFYEKSTPPNNLVNPVQELETERRV 189  
QY 243 KRKAPAPVLSPKTGVANTENTVSAGKDLSTSPKSPSPSPVLGKRPNAOSGLVWCKEVT 302  
DB 190 KRKAPAPVLSPKTGVANTENTVSAGKDLSTSPKSPSPSPVLGKRPNAOSGLVWCKEVT 249  
QY 303 KNYRGVKITNTTTSWRNGLSFCALIHPRPDLIDYKSLNPDDIKENKKAADGFAISIGIS 362  
DB 250 KNYRGVKITNTTTSWRNGLSFCALIHPRPDLIDYKSLNPDDIKENKKAADGFAISIGIS 309  
QY 363 RLLEPSDMVLAIIPDKLTVMTYLYOIRAHFSGQELANVOLEBNSKSTYKGVNEYETDINS 422  
DB 310 RLLEPSDMVLAIIPDKLTVMTYLYOIRAHFSGQELANVOLEBNSKSTYKGVNEYETDINS 369  
QY 423 SVDQEKFYAELSDLKREBELQOPISGAVDFLSQDVSFVNDGVSSESSEHQTPDDHLSR 482  
DB 370 SVDQEKFYAELSDLKREBELQOPISGAVDFLSQDVSFVNDGVSSESSEHQTPDDHLSR 429  
QY 483 STASPYCRATKSDTEPQKSOOSGRTSGSDPBGCSMTDSQOAVLLGKKRLKAFTLEL 542  
DB 430 STASPYCRATKSDTEPQKSOOSGRTSGSDPBGCSMTDSQOAVLLGKKRLKAFTLEL 489  
QY 543 SDLVVSDKKXDMSPFICEETDEBOKLTLDIGSNLEKLEKLENSRLBGRSPDESPIKTTS 602  
DB 490 SDLVVSDKKXDMSPFICEETDEBOKLTLDIGSNLEKLEKLENSRLBGRSPDESPIKTTS 549  
QY 603 LSPFSKLGYSYRDLDAKKKHAASLRQTESPDADRTTLNADHSSKIYQHRLLSRQBEL 662  
DB 550 LSPFSKLGYSYRDLDAKKKHAASLRQTESPDADRTTLNADHSSKIYQHRLLSRQBEL 609  
QY 663 KERARVLLBOARPAALKAGKHNNTATPPCNNOLSDQOEBRRROLREBAROLIBAR 722  
DB 610 KERARVLLBOARPAALKAGKHNNTATPPCNNOLSDQOEBRRROLREBAROLIBAR 669  
QY 723 SGVMSSELPSYGEMAAEKLKERSKASGDENDNIEIDTNEELPESFVVGGBELTNLENDL 782  
DB 670 SGVMSSELPSYGEMAAEKLKERSKASGDENDNIEIDTNEELPESFVVGGBELTNLENDL 729  
QY 783 DTEPQNSKLVLDLKLKLEVOPOVANSPPSSAAQAVTBESSFQDMKSGTEDLRTERLOKTT 842  
DB 730 DTEPQNSKLVLDLKLKLEVOPOVANSPPSSAAQAVTBESSFQDMKSGTEDLRTERLOKTT 789

QY 843 ERFNPVVPFSDSTVRKTQLOSFQYIENRPEMRQORSIOEDTKGNEKAATETORKP 902  
DB 790 ERFNPVVPFSDSTVRKTQLOSFQYIENRPEMRQORSIOEDTKGNEKAATETORKP 849  
QY 903 SEDEVLNKGFSDSYQYVVGELALENFOKQIDTPALVERKRLRYLMDTGRNTEEBEAMQ 962  
DB 850 SEDEVLNKGFSDSYQYVVGELALENFOKQIDTPALVERKRLRYLMDTGRNTEEBEAMQ 909  
QY 963 EWFVLVKKKVALPRNMQLSLEKEHDLERRYELTNLEBAMALIEDMQTEAQRRBQL 1022  
DB 910 EWFVLVKKKVALPRNMQLSLEKEHDLERRYELTNLEBAMALIEDMQTEAQRRBQL 969  
QY 1023 LLDLVLVKKRDALVBDLPAQEKQAEEDHELTERTLEONKGRKAKKBEKCVLQ 1076  
DB 970 LLDLVLVKKRDALVBDLPAQEKQAEEDHELTERTLEONKGRKAKKBEKCVLQ 1023  
RESULT 8  
ABG22883  
ID ABG22883 standard; protein; 1023 AA.  
XX  
AC ABG22883;  
XX  
DT 18-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #22874.  
XX  
KM Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KM food supplement; medical imaging; diagnostic; genetic disorder.  
XX Homo sapiens.  
XX  
OS WO200175067-A2.  
XX  
PN 11-OCT-2001.  
XX  
PD 30-MAR-2001; 2001WO-US008631.  
XX  
PF 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
XX (HYSB-) HYSBQ INC.  
XX  
PA Dmanac RT, Ldu C, Tang YT;  
XX  
PI WPI; 2001-639362/73.  
XX  
DR N-PSDB; AAS87070.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensic; gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
XX Claim 20; SEQ ID NO 53242; 103pp; English.  
XX  
PS The invention relates to isolated polynucleotide (I) and polypeptide (II)  
XX sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensic; gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic



CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp://wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 1023 AA;

Query Match 94.2%; Score 5173; DB 4; Length 1023;  
Best Local Similarity 99.7%; Pred. No. 1,6e-310;  
Matches 1011; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 63 IGNLDDFEEDNEDDENRVOEKAATTELINKLFLDEAKDLATVNSNPFDDPAAB 122  
DB 10 LANLDDFEEDNEDDENRVOEKAATTELINKLFLDEAKDLATVNSNPFDDPAAB 69  
QY 123 LNPFGDPSEBPIETETASPRKTEDESFYNSYNPFKEVQTPQYLNPFDEEAPVITKDSPP 182  
DB 70 LNPFGDPSEBPIETETASPRKTEDESFYNSYNPFKEVQTPQYLNPFDEEAPVITKDSPP 129  
QY 183 QSTGRKNIKRPVDMSKYLADSKTEEBELDESNPFYERKSTPPNNLVNPFQELTERRV 242  
DB 130 QSTGRKNIKRPVDMSKYLADSKTEEBELDESNPFYERKSTPPNNLVNPFQELTERRV 189  
QY 243 KRKAAPAPVLSPTKGVNMENTVSAGKDLSTSPKSPISPVYGRKPNASQSLVWCKEVT 302  
DB 190 KRKAAPAPVLSPTKGVNMENTVSAGKDLSTSPKSPISPVYGRKPNASQSLVWCKEVT 249  
QY 303 KNYGKVTNTTTSWRNGLSFCAIILHFRPDLIDYKSLNPQDICKNNKAYDGPASIGIS 362  
DB 250 KNYGKVTNTTTSWRNGLSFCAIILHFRPDLIDYKSLNPQDICKNNKAYDGPASIGIS 309  
QY 363 RLLEPSDVAFLAIPKLTVMYTYLQIRAFSGQELNVQIENSSEKSTYKVGNYETDINS 422  
DB 310 RLLEPSDVAFLAIPKLTVMYTYLQIRAFSGQELNVQIENSSEKSTYKVGNYETDINS 369  
QY 423 SVDEKFAEASDLKREBELQOPIGAVDFLSQDDSVVNSGVGESESEHQTPDDHLS 482  
DB 370 SVDEKFAEASDLKREBELQOPIGAVDFLSQDDSVVNSGVGESESEHQTPDDHLS 429  
QY 483 STASPYCRRKTSDETPQKSSQSGRTSGSDPQICSNSTDSYQAVTLGKGLKAKETEL 542  
DB 430 STASPYCRRKTSDETPQKSSQSGRTSGSDPQICSNSTDSYQAVTLGKGLKAKETEL 489  
QY 543 SDLVYSDKKDMSPFICEETDEBQLOFLDIGSNLEKEKLENSRSLKEGRSPESPRIKTS 602  
DB 490 SDLVYSDKKDMSPFICEETDEBQLOFLDIGSNLEKEKLENSRSLKEGRSPESPRIKTS 549  
QY 603 LSPTSKLGYSYRDLAKKKHASLRQTESPDADRTTLNHAHSSKIVQHRLLSROBEL 662  
DB 550 LSPTSKLGYSYRDLAKKKHASLRQTESPDADRTTLNHAHSSKIVQHRLLSROBEL 609  
QY 663 KERARVLLEQARRDALAKGNKHNNTATPFCNQLSDQDEBERRQLREBARQLIAEAR 722  
DB 610 KERARVLLEQARRDALAKGNKHNNTATPFCNQLSDQDEBERRQLREBARQLIAEAR 669  
QY 723 SGVMSSELPSYGEMAAERKLKERSKASGDENNIRIDTNEKIPEGVVGSGDELNTLENDL 782  
DB 670 SGVMSSELPSYGEMAAERKLKERSKASGDENNIRIDTNEKIPEGVVGSGDELNTLENDL 729  
QY 783 DTPESNKLVDLKLKGLLEVOPOVANSPSAAQKAVTSSSQDMKSGTEDRTERLQKTY 842  
DB 730 DTPESNKLVDLKLKGLLEVOPOVANSPSAAQKAVTSSSQDMKSGTEDRTERLQKTY 789  
QY 843 ERFNRPVVSFSDSYVRKTQLOSPSYIENRPEMRORSIOEDTKKGNKEKAIETQOKP 902  
DB 790 ERFNRPVVSFSDSYVRKTQLOSPSYIENRPEMRORSIOEDTKKGNKEKAIETQOKP 849  
QY 903 SEDVLNKGFDQTSQYVGHLLAALNEKQIDTAAVVERKLRYLMDGRTTERBEAMQ 962  
DB 850 SEDVLNKGFDQTSQYVGHLLAALNEKQIDTAAVVERKLRYLMDGRTTERBEAMQ 909  
QY 963 EFMFLVNKNALIRRMNQLSLLEKHDLERRYELLNRELRLAMLAIEDWQTEAQKRRQL 1022  
DB 910 EFMFLVNKNALIRRMNQLSLLEKHDLERRYELLNRELRLAMLAIEDWQTEAQKRRQL 966

DB 910 EFMFLVNKNALIRRMNQLSLLEKHDLERRYELLNRELRLAMLAIEDWQTEAQKRRQL 966  
QY 1023 LLDLVLVNNKRDALVRLDLOEKAEEDEHLETTLEONKGAKEKCVLQ 1076  
DB 970 LLDLVLVNNKRDALVRLDLOEKAEEDEHLETTLEONKGAKEKCVLQ 1023

RESULT 9  
ID AAM41108 standard; protein; 1023 AA.  
XX  
XX AAM41108;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 6039.

Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
XX 26-DEC-2000; 2000WO-US04263.  
XX  
PR 23-DEC-1989; 99US-00471275.  
XX  
PR 21-JAN-2000; 2000US-00488725.  
XX  
PR 25-APR-2000; 2000US-0052317.  
XX  
PR 20-JUN-2000; 2000US-00598042.  
XX  
PR 19-JUL-2000; 2000US-00620312.  
XX  
PR 03-AUG-2000; 2000US-00653450.  
XX  
PR 14-SEP-2000; 2000US-00662191.  
XX  
PR 19-OCT-2000; 2000US-00693036.  
XX  
PR 29-NOV-2000; 2000US-00727344.  
XX  
XX  
XX (HYSB-) HYSBQ INC.

XX Tang YT, Liu C, Aundri V, Chen R, Ma Y, Qian XB, Ren F, Wang D,  
XX Wang Y, Wang Z, Wehrman T, Xu C, Xue AD, Yang Y, Zhang J, Zhao QA,  
XX Zhou P, Goodrich R, Drmanac RT;  
XX WPI; 2001-442253/47.  
XX N-PSDB; AAI60264.  
XX  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders such  
XX as central nervous system injuries.  
XX  
XX  
XX Example 2; SEQ ID NO 6039; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and the  
XX encoded polypeptides (AAM8642-AA42213) with noctropic,  
XX immunosuppressant and cytostatic activity. The polynucleotides are useful  
XX in gene therapy. A composition containing a polypeptide or polynucleotide  
XX of the invention may be used to treat diseases of the peripheral nervous  
XX system, such as peripheral nervous injuries, peripheral neuropathy and  
XX localised neuropathies and central nervous system diseases, such as  
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
XX utilisation of the activities such as: Immune system suppression,  
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
XX assays for receptor activity, arthritis and inflammation, leukemias and  
XX C.N.S disorders. Note: The sequence data for this patent did not form  
XX part of the printed specification  
XX  
XX Sequence 1023 AA;

Query Match 93.9%; Score 5152; DB 4; Length 1023;  
 Best Local Similarity 99.3%; Pred. No. 3,1e-309;  
 Matches 1007; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```

QY 63 IGLNDDEEDNEDDENRVNQEKAKITTELINLNFLEDAEKDLATVNSNPFDDPAE 122
D 10 LANLDDFEEDNEDDENRVNQEKAKITTELINLNFLEDAEKDLATVNSNPFDDPAE 69
QY 123 LNPFGDDESEPIETETASPRRTEDSFYNNNSYNPFKEVQTPYLNPFDPEAFVITIKOSP 182
D 70 LNPFGDDESEPIETETASPRRTEDSFYNNNSYNPFKEVQTPYLNPFDPEAFVITIKOSP 129
QY 183 QSTRKKRIKRPVDMKTYLADSKTEBEELDESNPFEYSKSTPPPNLNLVNPQELETERRV 242
D 130 QSTRKKRIKRPVDMKTYLADSKTEBEELDESNPFEYSKSTPPPNLNLVNPQELETERRV 189
QY 243 KRKAAPPVLSPKTGVINENTVSAKDLSTSPKSPISPVLAGKRPNASQSLVWCKEVT 302
D 190 KRKAAPPVLSPKTGVINENTVSAKDLSTSPKSPISPVLAGKRPNASQSLVWCKEVT 249
QY 303 KNYGVKITNTTTSWRNGLSFCALIHRRPDLIDYKSLNPQDIKENNKKAVDGFASTIGIS 362
D 250 KNYGVKITNTTTSWRNGLSFCALIHRRPDLIDYKSLNPQDIKENNKKAVDGFASTIGIS 309
QY 363 RLLEPSDMVLTAIPDKLTVMTYLYQIRAHFSGQELANVQIENSSTKYKGVNETDPTS 422
D 310 RLLEPSDMVLTAIPDKLTVMTYLYQIRAHFSGQELANVQIENSSTKYKGVNETDPTS 369
QY 423 SVDEKRYAELSDIKREBELQOPIGAVDFLSQDDSVFVNDGVESESEHQTPDDHLS 482
D 370 SVDEKRYAELSDIKREBELQOPIGAVDFLSQDDSVFVNDGVESESEHQTPDDHLS 429
QY 483 STASPYGRTKSDTEPOKSOOSSGRTSGDDPGICSMTDSYQAOVLLGKRLLAETLEL 542
D 430 STASPYGRTKSDTEPOKSOOSSGRTSGDDPGICSMTDSYQAOVLLGKRLLAETLEL 489
QY 543 SDLYVSDKKDMSPFICEETDEQKLOTLDIGSNLEKEKLENSSLSECRSPESPICKTS 602
D 490 SDLYVSDKKDMSPFICEETDEQKLOTLDIGSNLEKEKLENSSLSECRSPESPICKTS 549
QY 603 LSPTSKLGYSYRDLDAKKKHAASLRQTESDDPDAKRTLLNADHSSKIVQHRLLSROBEL 662
D 550 LSPTSKLGYSYRDLDAKKKHAASLRQTESDDPDAKRTLLNADHSSKIVQHRLLSROBEL 609
QY 663 KEPRVULLEQARRRAALKAGNKHNTNTATPCNPNQLSQDOEBERRQLREARQLIAR 722
D 610 KEPRVULLEQARRRAALKAGNKHNTNTATPCNPNQLSQDOEBERRQLREARQLIAR 669
QY 723 SGVVMSELPSYGEWAAREKLEKRSKASGENDNIEIDTNEEIPBGFVVGSGDELNTLENDL 782
D 670 SGVVMSELPSYGEWAAREKLEKRSKASGENDNIEIDTNEEIPBGFVVGSGDELNTLENDL 729
QY 783 DTPEQNSKLVDLKIKLLEVOQVANSPPSSAAQAVTESSBQDKSGTEDLRTERLQKTT 842
D 730 DTPEQNSKLVDLKIKLLEVOQVANSPPSSAAQAVTESSBQDKSGTEDLRTERLQKTT 789
QY 843 ERRRNPVVFESKDSVVRKTQULSFQYIENRPEMKRQSIQBDTKKNEBKRAITETQKRP 902
D 790 ERRRNPVVFESKDSVVRKTQULSFQYIENRPEMKRQSIQBDTKKNEBKRAITETQKRP 849
QY 903 SEDEVLNKGFQDTSQYVVGELAALENBOKQIDTAAALVEKKRLRYLMDTGRVTEEBAMQ 962
D 850 SEDEVLNKGFQDTSQYVVGELAALENBOKQIDTAAALVEKKRLRYLMDTGRVTEEBAMQ 909
QY 963 EFWPLVANKKALIRRMNQSLLEKREHDLERRYELLNLEBLRMLALIEDMQKTEAQKREQL 1022
D 910 EFWPLVANKKALIRRMNQSLLEKREHDLERRYELLNLEBLRMLALIEDMQKTEAQKREQL 969
QY 1023 LLEBELVALVKKRDALVNDLDAQEQAEEBDEHLERTLEQNKKGKRAKKEKCVLQ 1076
D 970 LLEBELVALVKKRDALVNDLDAQEQAEEBDEHLERTLEQNKKGKRAKKEKCVLQ 1023

```

RESULT 10  
 AEA21081  
 ID AEA21081 standard; protein: 1023 AA.  
 XX  
 AC AEA21081;  
 XX  
 DT 11-AUG-2005 (first entry)  
 XX  
 DE Novel human polypeptide SEQ ID NO 1775.  
 XX  
 XX vulnerable; CNS-gen.; gene therapy; diagnostic; forensic; mapping;  
 KM DNA purification; protein purification; osteoarthritis; antiarthritic;  
 KM osteopathic; musculoskeletal disease; osteoporosis; endocrine disease;  
 KM periodontal disease; antiinflammatory; mouth disease; burns; injury;  
 KM peripheral neuropathy; Alzheimer's disease; neuroprotective; noctropic;  
 KM degenerative; Parkinson's disease; antiparkinsonian; neurological disease;  
 KM cerebrovascular ischemia; cerebroprotective; vasotropic;  
 KM cardiovascular disease; autoimmune disease; immunosuppressive;  
 KM immune disorder; viral infection; virucide; infection; cancer;  
 KM cytostatic; neoplasm.  
 KM  
 XX Homo sapiens.  
 OS  
 PN W02005049806-A2.  
 XX  
 PD 02-JUN-2005.  
 XX  
 PF 11-MAR-2004; 2004MO-US007412.  
 XX  
 PR 14-MAR-2003; 2003US-00389559.  
 XX  
 PA (NUVE-) NUVELO INC.  
 XX  
 PI Tang TY, Wang J, Wang ZM, Zhang J, Ren F, Zhou P, Ma Y,  
 PI Ghosh M, Xue A, Asundi V, Zhao Q, Wang D, Goodrich R, Chen R,  
 PI Wehrman T, Wang G, Boyle B;  
 XX  
 DR WPI; 2005-417730/42.  
 XX  
 XX New polynucleotide encoding a polypeptide with biological activity,  
 PT useful for treating a disease or disorder, e.g. osteoarthritis, burns,  
 PT CNS and peripheral disease, stroke, autoimmune disorders, viral  
 PT infection, or cancer.  
 XX  
 PS Example 3; SEQ ID NO 1775; 500bp; English.  
 XX  
 CC The invention describes a new isolated polynucleotide (1) encoding a  
 CC polypeptide with biological activity comprising: a nucleotide sequence of  
 CC SEQ ID NOS: 1-567 (fully defined); a nucleotide sequence that hybridizes  
 CC to the sequence of (1) under stringent hybridization conditions; or a  
 CC nucleotide sequence having greater than 9% sequence identity with the  
 CC sequence of (1). Also described are: a(n) (expression)vector comprising  
 CC (1); a host cell genetically engineered to comprise (1) operatively,  
 CC associated with a regulatory sequence that modulates expression of the  
 CC polynucleotide in the host cell; an isolated polypeptide comprising a  
 CC sequence of SEQ ID NOS: 568-1134 (fully defined) where the polypeptide  
 CC is a polypeptide encoded by (1); or a polypeptide encoded by a  
 CC polynucleotide hybridizing under stringent conditions with any one of SEQ  
 CC ID NOS: 1-567; a composition comprising the polypeptide of (3) and a  
 CC carrier; an antibody directed against the polypeptide of (3); a method  
 CC for detecting (1) in a sample; a method for detecting the polypeptide of  
 CC (3) in a sample; a method for identifying a compound that binds to the  
 CC polypeptide of (3); a method of producing the polypeptide of (3); and a  
 CC collection of polynucleotides, where the collection comprising of at  
 CC least one of SEQ ID NOS: 1-567. (1) is a polynucleotide comprising any of  
 CC the sequences of SEQ ID NOS: 1-567 encoding a polypeptide with biological  
 CC activity, which comprises any of the amino acid sequence of SEQ ID NOS:  
 CC 568-1134. All sequences are fully defined in the specification. The  
 CC sequences and methods are useful in diagnostics, forensic, and gene  
 CC mapping, in identifying of mutations responsible for genetic disorders or  
 CC other traits, in assessing biodiversity, and for producing many other  
 CC types of data and products dependent on DNA and amino acid sequences. The







CC encoded polypeptides (AAM3642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localized neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
CC part of the printed specification

**SQ Sequence 996 AA;**

Query Match	91.7%	Score 5033	DB 4	Length 996
-------------	-------	------------	------	------------

Best Local Similarity 99.3%; Pred. No. 6.7e-302;  
Matches 985; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy	85	IKAAKITELINGLNFDEAEKDLATVNSNPDDPDAELNPPGDPSSEPIETETASPRKT	144
Db	5	KSASSSEBELINKNLFLDEAEKDLATVNSNPDDPDAELNPPGDPSSEPIETETASPRKT	64
Qy	145	EDSFNNSYNPEKEVOTPOLYLNPDEPEAFVITIKOSPPOSTRKXNIRPVMDSKYLADSS	204
Db	65	EDSFNNSYNPEKEVOTPOLYLNPDEPEAFVITIKOSPPOSTRKXNIRPVMDSKYLADSS	124
Qy	205	KTEBEELDESNPFFIYPPKSTPPPNNTLVNPQELETERRVKRAKAPAPVLSPKTVLMENTV	264
Db	125	KTEBEELDESNPFFIYPPKSTPPPNNTLVNPQELETERRVKRAKAPAPVLSPKTVLMENTV	184
Qy	265	SAGKDLSTSPKPSPIPSPIGKRKPNASOSLWMCCKEVTGNXGAKTNTFTSGRNLSTSC	324
Db	185	SAGKDLSTSPKPSPIPSPIGKRKPNASOSLWMCCKEVTGNXGAKTNTFTSGRNLSTSC	244
Qy	325	AILHHFRDLIDYKSLNPDIKENNKKAVDGPASIGISRLLEPSDMVLAIPDKLVVMY	384
Db	245	AILHHFRDLIDYKSLNPDIKENNKKAVDGPASIGISRLLEPSDMVLAIPDKLVVMY	304
Qy	385	LYOIRAHSGGELNVAQIENSCKSTYKXGNYETDNTSSVDOKFYABLSDLKREBELQO	444
Db	305	LYOIRAHSGGELNVAQIENSCKSTYKXGNYETDNTSSVDOKFYABLSDLKREBELQO	364
Qy	445	PISGAVNDLSODDSVFVNDSGVGESESEHQTDHDLSPETASPYCGRTSDDTPEKSOOS	504
Db	365	PISGAVNDLSODDSVFVNDSGVGESESEHQTDHDLSPETASPYCGRTSDDTPEKSOOS	424
Qy	505	SGRTSGSDPDGICSNWTBQAVLLGKKRLLAETLELSDLVYSDDKDKMSPPFICEETD	564
Db	425	SGRTSGSDPDGICSNWTBQAVLLGKKRLLAETLELSDLVYSDDKDKMSPPFICEETD	484
Qy	565	EOKLQTLDIGSNLEKLEKLENSKSLCERSDPESPICKTSLISPTSKLGYSRDLDAKKKH	624
Db	485	EOKLQTLDIGSNLEKLEKLENSKSLCERSDPESPICKTSLISPTSKLGYSRDLDAKKKH	544
Qy	625	ASLAEQTESDDPADRTLLNHAHDSSKTVQHRLLSROBELKERARVYLLEQARRDALAKGK	684
Db	545	ASLAEQTESDDPADRTLLNHAHDSSKTVQHRLLSROBELKERARVYLLEQARRDALAKGK	604
Qy	685	HNNTATATFCNROLSDQODEERRRQDLREBARQDLIENASGVGMSELPSYGBMAAEKLKER	744
Db	605	HNNTATATFCNROLSDQODEERRRQDLREBARQDLIENASGVGMSELPSYGBMAAEKLKER	664
Qy	745	SKASGENDENIETIDNBEIPIRGVVGGBDELTNLENDLTPPEONSKLYDLKLKULLEVOP	804
Db	665	SKASGENDENIETIDNBEIPIRGVVGGBDELTNLENDLTPPEONSKLYDLKLKULLEVOP	724
Qy	805	QVANSPPSAQAKVATESSEQDMKSGTEDLRTERLOKTTERRPNPVVFSKOSTYVRKTQLOS	864
Db	725	QVANSPPSAQAKVATESSEQDMKSGTEDLRTERLOKTTERRPNPVVFSKOSTYVRKTQLOS	784

QY	865	SSQYIENPBNKROSI	IEDPTKKGNEBKAAL	TTETORXSDEBVLNKGKEDTSQYVVGELA	924
Db	785	FSQYIENPBNKROSI	IEDPTKKGNEBKAAL	TTETORXSDEBVLNKGKEDTSQYVVGELA	844
QY	925	ALENEQOQIDTRAL	VEKRLRYLMDTGNTSEEBAMQEMFELV	NKQKALIRBNQISLL	984
Db	845	ALENEQOQIDTRAL	VEKRLRYLMDTGNTSEEBAMQEMFELV	NKQKALIRBNQISLL	904
QY	985	EKEHDLERRYELN	RELPAMLAIEDMQTEBAQKREQLL	DELVALVYNKDALVFDLDAQ	1044
Db	905	EKEHDLERRYELN	RELPAMLAIEDMQTEBAQKREQLL	DELVALVYNKDALVFDLDAQ	964
QY	1045	EKOABEEDHEHLER	TLEONKKGMAKKEECVIAQ	1076	
Db	965	EKOABEEDHEHLER	TLEONKKGMAKKEECVIAQ	996	

RESULT 14  
ABA20269

AC ABA20269;

DT 11-AUG-2005 (first entry)

DB Novel human polypeptide SEQ ID NO 963.

KM vulnérability; CNS-gen.; gene therapy; diagnostic; forensic; mapping;  
KM DNA purification; protein purification; osteoarthritis; antiarthritis;  
KM osteopathic; musculoskeletal disease; osteoporosis; endocrine disease;  
KM periodontal disease; antiinflammatory; mouth disease; burns; injury;  
KM peripheral neuropathy; Alzheimer disease; neuroprotective; nootropic;  
KM degenerative; parkinsons disease; antiparkinsonian; neurological disease;  
KM cerebrovascular ischemia; cerebroprotective; vasotropic;  
KM cardiovascular disease; autoimmune disease; immunosuppressive;  
KM immune disorder; viral infection; viroicide; infection; cancer;  
KM cytostatic; neoplasm.

OS Homo sapiens

PN WO2005049806-A2

PD 02-JUN-2005

PF 11-MAR-2004; 2004WO-US007412.

PR 14-MAR-2003; 2003US-00389559.

PA (NUVE-) NUVEIO INC.

PI	Tang TY, Wang J, Wang ZW, Zhang J, Ren F, Zhou P, Ma Y;
PI	Ghosh M, Xue A, Asundi V, Zhao Q, Wang D, Goodrich R, Chen R;
PI	Wehrman T, Weng G, Boyle B;

DR WPI; 2005-417730/42.  
DR N-PSDB; AEA19702.

PT New polynucleotide encoding a polypeptide with biological activity,  
PT useful for treating a disease or disorder, e.g. osteoarthritis, burns  
PT CNS and peripheral disease, stroke, autoimmune disorders, viral  
PT infection, or cancer.

PS Claim 20; SEQ ID NO 963; 500pp; English.

CC The invention describes a new isolated polynucleotide (I) encoding a  
CC polypeptide with biological activity comprising: a nucleotide sequence of  
CC SEQ ID NOS: 1-567 (fully defined); a nucleotide sequence that hybridizes  
CC to the sequence of (I) under stringent hybridization conditions; or a  
CC nucleotide sequence having greater than 99% sequence identity with the  
CC sequence of (I). Also described are: a(n) (expression) vector comprising  
CC (I); a host cell genetically engineered to comprise (I) operatively,  
CC associated with a regulatory sequence that modulates expression of the  
CC polynucleotide in the host cell; an isolated polypeptide comprising a



CC sequence of SEQ ID NOS: 568-1134 (fully defined), where the polypeptide  
 CC is a polypeptide encoded by (1); or a polypeptide encoded by a  
 CC polynucleotide hybridizing under stringent conditions with any one of SEQ  
 CC ID NOS: 1-567; a composition comprising the polypeptide of (3) and a  
 CC carrier; an antibody directed against the polypeptide of (3); a method  
 CC for detecting (1) in a sample; a method for detecting the polypeptide of  
 CC (3) in a sample; a method for identifying a compound that binds to the  
 CC polypeptide of (3); a method for producing the polypeptide of (3); and a  
 CC collection of polynucleotides, where the collection comprising of at  
 CC least one of SEQ ID NOS: 1-567. (1) is a polynucleotide comprising any of  
 CC the sequences of SEQ ID NOS: 1-567 encoding a polypeptide with biological  
 CC activity, which comprises any of the amino acid sequence of SEQ ID NOS:  
 CC 568-1134. All sequences are fully defined in the specification. The  
 CC sequences and methods are useful in diagnostics, forensic, and gene  
 CC mapping, in identifying of mutations responsible for genetic disorders or  
 CC other traits, in assessing biodiversity, and for producing many other  
 CC types of data and products dependent on DNA and amino acid sequences. The  
 CC composition and method are useful for treating a disease or disorder,  
 CC e.g., osteoporosis, osteoarthritis, periodontal disease, burns, CNS and  
 CC peripheral disease, Alzheimer's disease, Parkinson's disease, stroke,  
 CC autoimmune disorders, viral infection, or cancer. This is the amino acid  
 CC sequence of a novel polypeptide of the invention.

XX Sequence 960 AA;

Query Match 87.8%; Score 4817; DB 9; Length 960;

Best Local Similarity 95.8%; Pred. No. 1.4e-288; Indels 36; Gaps 1;

Matches 950; Conservative 2; Mismatches 4; Indels 36; Gaps 1;

QY 85 EKAAKTELINLNFLEBAEKDLATVNSNPPDDPAALNPPGDDSEPIETETASPRKT 144  
 DB 5 KSASSESELINKLNFLEBAEKDLATVNSNPPDDPAALNPPGDDSEPIETETASPRKT 64  
 QY 145 EDSFNNSYNPKKEVQTPQYLNPPDEPAFVITIKOSPSTGRKRIKIRPYMSKTIYADSS 204  
 DB 65 BDFSNNSYNPKKEVQTPQYLNPPDEPAFVITIKOSPSTGRKRIKIRPYMSKTIYADSS 124  
 QY 205 KTEBELLESNPFYKSTPPNNLVNPPQELERVRVKAAPAPVLSPTKGVNENTV 264  
 DB 125 KTEBELLESNPFYKSTPPNNLVNPPQELERVRVKAAPAPVLSPTKGVNENTV 184  
 QY 265 SAGKDLSTSPKSPISPVLAGRKPNASQSLVWCKEVTKNYKGVKITNFTTSMRGLSFC 324  
 DB 185 SAGKDLSTSPKSPISPVLAGRKPNASQSLVWCKEVTKNYKGVKITNFTTSMRGLSFC 244  
 QY 325 AILHFRPDLLDYKSLNPPDIKNNKAYDGPASIGISRLLEPSPMVLAIIPDKITVMTY 384  
 DB 245 AILHFRPDLLDYKSLNPPDIKNNKAYDGPASIGISRLLEPSPMVLAIIPDKITVMTY 304  
 QY 385 LYQIRAHSGOELNVVOJBENSSKSTYKGVNFTDNTSSVQEKFYAELSDKREPELQ 444  
 DB 305 LYQIRAHSGOELNVVOJBENSSKSTYKGVNFTDNTSSVQEKFYAELSDKREPELQ 364  
 QY 445 PISGAVDFLSQDDSVFVNDGVSSESSEHQTPDDHLSPTASPCRTKSTDEPOKSQOS 504  
 DB 365 PISGAVDFLSQDDSVFVNDGVSSESSEHQTPDDHLSPTASPCRTKSTDEPOKSQOS 424  
 QY 505 SGRISGSDDPGICSNVTSTQAVLLGKRLKKAFTLEISDLVYSDKKDMSPPICBEETD 564  
 DB 425 SGRISGSDDPGICSNVTSTQAVLLGKRLKKAFTLEISDLVYSDKKDMSPPICBEETD 484  
 QY 565 EOKIQTLLIGSLTEKLENSRSLERCDSPSPICKTSLSPSKIGYSYRDLAKK 624  
 DB 485 EOKIQTLLIGSLTEKLENSRSLERCDSPSPICKTSLSPSKIGYSYRDLAKK 544  
 QY 625 ASLRQTESDPAADRTTLNHAHDSKIVQIRLLSRQBELKERAVLLLEQARRDALAKGK 684  
 DB 545 ASLRQTESDPAADRTTLNHAHDSKIVQIRLLSRQBELKERAVLLLEQARRDALAKGK 604  
 QY 685 HNTNATPFCNKRQLSDQDEERRQLREBARQLIABASGVMSSELPSYGEVMAAKLER 744  
 DB 605 HNTNATPFCNKRQLSDQDEERRQLREBARQLIABASGVMSSELPSYGEVMAAKLER 664

QY 745 SKASGENDNIEIDTNERIPGPFVVGDEBLTNLENDLDTPEONSKLVDLKLLLEVOP 804  
 DB 665 SKASG-----EQNKVLVDLKKLLLEVOP 688  
 QY 805 QVANSPPSAQAVTESSEBOMKSTBULRTBRLOXTERRRPNPVFSKSTVKTQLOS 864  
 DB 689 QVANSPPSAQAVTESSEBOMKSTBULRTBRLOXTERRRPNPVFSKSTVKTQLOS 748  
 QY 865 FSOYIENPKMKRORSIOEDTKNGEKAATTEORKESEDEVINKGFKDTSQYVVGELA 924  
 DB 749 FSOYIENPKMKRORSIOEDTKNGEKAATTEORKESEDEVINKGFKDTSQYVVGELA 808  
 QY 925 ALENEKOQIDTRALVVERKRLYLMDTGNTEEBEAMQEMFVLVNNKVALIRNNQLSTL 984  
 DB 809 ALENEKOQIDTRALVVERKRLYLMDTGNTEEBEAMQEMFVLVNNKVALIRNNQLSTL 868  
 QY 985 EKHDLERRYLNLRELPMALIEBDMQTEBAQKRREQLLDBLVALVNRDALVDDLDQ 1044  
 DB 869 EKHDLERRYLNLRELPMALIEBDMQTEBAQKRREQLLDBLVALVNRDALVDDLDQ 928  
 QY 1045 EKQAEERDEHLERTLEONKGMKAKKEKCVIQ 1076  
 DB 929 EKQAEERDEHLERTLEONKGMKAKKEKCVIQ 960

RESULT 15

ADU70136  
 ID ADU70136 standard; protein; 882 AA.

AC ADU70136;

DT 06-MAY-2004 (first entry)

DE Human heat mitochondrial protein as a therapeutic target SegID1942.

KM mitochondrial; human; screening assay; diabetes mellitus;  
 KM Huntington's disease; osteoarthritis;

KM Leber's hereditary optic neuropathy; LHON;  
 KM mitochondrial encephalopathy lactic acidosis and stroke; MELAS;

KM myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;  
 KM neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;  
 KM osteopathic; ophthalmological; cyostatic.

OS Homo sapiens.

FN WO2003087768-A2.

PD 23-OCT-2003.

PF 04-APR-2003; 2003WO-US010870.

PR 12-APR-2002; 2002US-0372843P.

PR 17-JUN-2002; 2002US-0389987P.

PR 20-SEP-2002; 2002US-0412418P.

PA (MITO-) MITOKOR.

PI (BUCK-) BUCK INST AGE RES.

PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;  
 PI Warmack DG;

XX WPI; 2003-845369/78.

PT Identifying a mitochondrial target for drug screening assays and for  
 PT treating diseases associated with altered mitochondrial function,  
 PT comprises detecting a modified polypeptide in a sample and correlating  
 PT with the disease.

PT Claim 1; SEQ ID NO 1942; 180pp; English.

PS This invention relates to novel mitochondrial targets that can be used  
 PS for therapeutic intervention in treating a disease associated with  
 PS altered mitochondrial function. Specifically, it refers to a method for

CC altered mitochondrial function. Specifically, it refers to a method for

CC altered mitochondrial function. Specifically, it refers to a method for

CC Identifying proteins of the human heart mitochondrial proteome that are  
 CC useful for drug screening assays, as well as therapeutic targets. The  
 CC present invention describes a method for identifying such proteins. The  
 CC can be used in the treatment of various diseases associated with altered  
 CC mitochondrial function including diabetes mellitus, Huntington's disease,  
 CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial  
 CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy  
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these  
 CC compositions have neuroprotective, neurotropic, antidiabetic,  
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and  
 CC cyostatic activities. This polypeptide sequence is a human heart  
 CC mitochondrial protein of the invention.  
 XX  
 SQ Sequence 882 AA:

Query Match 81.7%; Score 4485; DB 7; Length 882;  
 Best Local Similarity 100.0%; Pred. No. 4e-268;  
 Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 MSKTLVADSSKTEEBELDESNPFPKSTPPNNLVNVOELFETRRYKRAPVPLSP 254  
 DB 1 MSKTLVADSSKTEEBELDESNPFPKSTPPNNLVNVOELFETRRYKRAPVPLSP 60  
 QY 255 KTVGLNENTVSAAGDLSPPKSPPISEVLAGRKPNASQSLVWCSEVTKYRGVKTNFT 314  
 DB 61 KTVGLNENTVSAAGDLSPPKSPPISEVLAGRKPNASQSLVWCSEVTKYRGVKTNFT 120  
 QY 315 TSWNGSLSPCALIHHFRPDLIDYKSLNPQDIKENNKAYDGFASIGISRLPESDMVLLA 374  
 DB 121 TSWNGSLSPCALIHHFRPDLIDYKSLNPQDIKENNKAYDGFASIGISRLPESDMVLLA 180  
 QY 375 IPDLATVMTYLYOIRAHPSGQELNVVOEENSSTKYKVGNETDTNSVDQEKFYAELS 434  
 DB 181 IPDLATVMTYLYOIRAHPSGQELNVVOEENSSTKYKVGNETDTNSVDQEKFYAELS 240  
 QY 435 DLKKEPELQPIISGAVDFLSQDSDVFNVDGVSSESEHQTPDDHLSPTASPYCRRTKS 494  
 DB 241 DLKKEPELQPIISGAVDFLSQDSDVFNVDGVSSESEHQTPDDHLSPTASPYCRRTKS 300  
 QY 495 DTEPQKQSSGRTSGSDDDPGICNTDSTQAOVLGKKRLKATLELSDLVSDKKDM 554  
 DB 301 DTEPQKQSSGRTSGSDDDPGICNTDSTQAOVLGKKRLKATLELSDLVSDKKDM 360  
 QY 555 SPPPICETBEQKQOTLDIGSNLEKELENSRSLCSDSPESPIKKTSLSPSTGLGYSYS 614  
 DB 361 SPPPICETBEQKQOTLDIGSNLEKELENSRSLCSDSPESPIKKTSLSPSTGLGYSYS 420  
 QY 615 RDLDLAKKQASLQTESDDADRTTLNHADHSSKIYQHRLLSRQBELKERAVYLEQAR 674  
 DB 421 RDLDLAKKQASLQTESDDADRTTLNHADHSSKIYQHRLLSRQBELKERAVYLEQAR 480  
 QY 675 RDAALKAQGNNTATPFCRQSLSDQODEERRQLRBRARQLIAEARSQVMSSELPYSG 734  
 DB 481 RDAALKAQGNNTATPFCRQSLSDQODEERRQLRBRARQLIAEARSQVMSSELPYSG 540  
 QY 735 EMAAEKLERSKASGDENDNIEIDTNEIPEGFVVGDDDELTNLENDLTPQNSKLVDL 794  
 DB 541 EMAAEKLERSKASGDENDNIEIDTNEIPEGFVVGDDDELTNLENDLTPQNSKLVDL 600  
 QY 795 KLKKLLEVOQOVANSPSSAAQKAVTESSEDMKSGTEDLRTERLOKTERFRNPFVSKD 854  
 DB 601 KLKKLLEVOQOVANSPSSAAQKAVTESSEDMKSGTEDLRTERLOKTERFRNPFVSKD 660  
 QY 855 STVRKTLQSPQYIENRPEMKRQSTIQEDTKKNEKAAITETORXPSDEVLANKGPKD 914  
 DB 661 STVRKTLQSPQYIENRPEMKRQSTIQEDTKKNEKAAITETORXPSDEVLANKGPKD 720  
 QY 915 TSOYVVELALAEQEQIDTRALVEKRLRYLMDTGRNTEEBEAMQEWPMLVNKKNAL 974  
 DB 721 TSOYVVELALAEQEQIDTRALVEKRLRYLMDTGRNTEEBEAMQEWPMLVNKKNAL 780  
 QY 975 IRRNQSLLEKHEHLERRYELLNRELRLAMLAIEDMQTEAQKRQQLLDELVALVNR 1034  
 DB 781 IRRNQSLLEKHEHLERRYELLNRELRLAMLAIEDMQTEAQKRQQLLDELVALVNR 840

DB 781 IRRNQSLLEKHEHLERRYELLNRELRLAMLAIEDMQTEAQKRQQLLDELVALVNR 840  
 QY 1035 DALVRDLDAAQEQAEEDDEHLEERTLEONKGMKKEKCVLQ 1076  
 DB 841 DALVRDLDAAQEQAEEDDEHLEERTLEONKGMKKEKCVLQ 882  
 Search completed: May 20, 2006, 17:09:12  
 Job time : 207 secs



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using SW model

Run on: May 20, 2006, 17:14:58 ; Search time 53 Seconds  
(without alignments)  
1777.037 Million cell updates/sec

Title: US-09-856-723a-8

Perfect score: 5489

Sequence: 1 MKQYASPMPTOTDVLKFKFP.....RTLEQNGKMKAKKEKCVLQ 1076

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_AA:\*

1: /EMC\_Celestra\_SIDS3/ptodata/2/1aa/5.COMB.pep:\*

2: /EMC\_Celestra\_SIDS3/ptodata/2/1aa/6.COMB.pep:\*

3: /EMC\_Celestra\_SIDS3/ptodata/2/1aa/7.COMB.pep:\*

4: /EMC\_Celestra\_SIDS3/ptodata/2/1aa/H.COMB.pep:\*

5: /EMC\_Celestra\_SIDS3/ptodata/2/1aa/PTUS.COMB.pep:\*

6: /EMC\_Celestra\_SIDS3/ptodata/2/1aa/RB.COMB.pep:\*

7: /EMC\_Celestra\_SIDS3/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5486	99.9	1076	US-09-976-594-889	Sequence 889, App
2	326.5	5.9	3433	US-09-538-092-1136	Sequence 1136, Ap
3	307.5	5.6	2141	US-09-949-016-10918	Sequence 10918, A
4	301.5	5.5	3433	US-09-091-501B-10	Sequence 10, Appl
5	296	5.4	61	US-09-513-999C-7173	Sequence 7173, Ap
6	291.5	5.3	2154	US-08-841-349-4	Sequence 4, Appl
7	291.5	5.3	2154	US-09-431-184A-4	Sequence 4, Appl
8	289.5	5.3	2364	US-09-538-092-1243	Sequence 1243, Ap
9	286	5.2	2388	US-09-695-795A-2	Sequence 2, Appl
10	285	5.2	1026	US-09-949-016-6777	Sequence 6777, Ap
11	285	5.2	1034	US-09-949-016-10870	Sequence 10870, A
12	283.5	5.2	153	US-09-270-767-42273	Sequence 42273, A
13	276.5	5.0	912	US-09-664-958-8	Sequence 8, Appl
14	274	5.0	883	US-09-664-958-10	Sequence 10, Appl
15	270.5	4.9	2008	US-09-091-501B-8	Sequence 8, Appl
16	270.5	4.9	2186	US-09-949-016-10828	Sequence 10828, A
17	270.5	4.9	2349	US-09-538-092-914	Sequence 914, App
18	264	4.8	1078	US-09-248-796A-20284	Sequence 20284, A
19	264	4.8	2662	US-09-595-684B-31	Sequence 31, Appl
20	262.5	4.8	2663	US-09-538-092-1252	Sequence 1252, Ap
21	261.5	4.8	2375	US-09-538-092-1131	Sequence 1131, Ap
22	261	4.8	630	US-09-248-796A-20275	Sequence 20275, A
23	258.5	4.7	1093	US-09-949-016-11535	Sequence 11535, A
24	258.5	4.7	1093	PCT-US93-03077-1	Sequence 1, Appl
25	256	4.7	1976	US-09-538-092-1078	Sequence 1078, Ap
26	255	4.6	10182	US-09-134-001C-3159	Sequence 3159, Ap

27	254.5	4.6	3666	US-09-134-001C-5080	Sequence 5080, Ap
28	254	4.6	2600	US-09-949-016-7309	Sequence 7309, Ap
29	254	4.6	5024	US-09-710-279-2964	Sequence 2964, Ap
30	251	4.6	1588	PCT-US93-07261-11	Sequence 11, Appl
31	251	4.6	1663	PCT-US93-07261-16	Sequence 16, Appl
32	250.5	4.6	1960	US-09-949-016-10872	Sequence 10872, A
33	250	4.6	905	US-09-248-796A-16333	Sequence 16333, A
34	250	4.6	2733	US-09-949-016-11433	Sequence 11433, A
35	250	4.6	3259	US-09-949-016-6507	Sequence 6507, Ap
36	249.5	4.5	1960	US-09-538-092-1077	Sequence 1077, Ap
37	248	4.5	1958	US-10-028-946-4	Sequence 4, Appl
38	248	4.5	2053	US-09-964-956-11	Sequence 11, Appl
39	248	4.5	2054	US-10-028-946-2	Sequence 2, Appl
40	248	4.5	2066	US-09-964-956-9	Sequence 9, Appl
41	247.5	4.5	3878	US-09-914-259-11	Sequence 11, Appl
42	247	4.5	1972	US-08-875-435B-4	Sequence 4, Appl
43	245.5	4.5	1219	US-09-344-624-4	Sequence 4, Appl
44	245	4.5	924	US-09-248-796A-18798	Sequence 18798, A
45	243	4.4	1307	US-09-949-016-7561	Sequence 7561, Ap

ALIGNMENTS

RESULT 1					
US-09-976-594-889					
Sequence 889, Application US/09976594					
Patent No. 6673549					
GENERAL INFORMATION:					
APPLICANT: Furness, Michael					
APPLICANT: Buchbinder, Jenny					
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS					
FILE REFERENCE: PA-0041 US					
CURRENT APPLICATION NUMBER: US/09/976,594					
CURRENT FILING DATE: 2001-10-12					
PRIOR APPLICATION NUMBER: 60/240,409					
PRIOR FILING DATE: 2000-10-12					
NUMBER OF SEQ ID NOS: 1143					
SOFTWARE: PERL Program					
SEQ ID NO 889					
LENGTH: 1076					
TYPE: PRT					
ORGANISM: Homo sapiens					
FEATURE:					
NAME/KEY: misc_feature					
OTHER INFORMATION: Incyte ID No. 6673549 5202390CD1					
US-09-976-594-889					
Query Match					
Best Local Similarity 99.9%; Score 5486; DB 2; Length 1076;					
Matches 1075; Conservative 1; Mismatches 0; Indels 0; Gaps 0;					
QY	1	MKQYASPMPTOTDVLKFKPLSKKYVSAALQFSLSCTFLRGGKATDEDMQSLASVSMKQ	60		
DB	1	MKQYASPMPTOTDVLKFKPLSKKYVSAALQFSLSCTFLRGGKATDEDMQSLASVSMKQ	60		
QY	61	ADIGNLDDFEEDNEDDENRYNOEBAKATTELINKNLFLEBAEKDLATVNSNPPDDPA	120		
DB	61	ADIGNLDDFEEDNEDDENRYNOEBAKATTELINKNLFLEBAEKDLATVNSNPPDDPA	120		
QY	121	AEINPFGDPDEEPTTETASPRKTEDESYNNSYNPFXVOTPOYLNPDEBEAFVTTIDS	180		
DB	121	AEINPFGDPDEEPTTETASPRKTEDESYNNSYNPFXVOTPOYLNPDEBEAFVTTIDS	180		
QY	181	PPQSTRKKNIRPVMSKTYLADSSKTEBBELDESNPFYEPKSTPPNNLVNPFVDELTER	240		
DB	181	PPQSTRKKNIRPVMSKTYLADSSKTEBBELDESNPFYEPKSTPPNNLVNPFVDELTER	240		
QY	241	RVKKAPAPPVLSKRTGYLMENTYSAGKDLSTSPKSPRIPSPVGRKPNASQSLVWCKE	300		
DB	241	RVKKAPAPPVLSKRTGYLMENTYSAGKDLSTSPKSPRIPSPVGRKPNASQSLVWCKE	300		
QY	301	VTKNYRGVKTNPFTTSWRNGLSFCALLHHFRPDLIDYKSLNPQIKENNKAYDGFASIG	360		

```
Db 301 VTKAVRGVLTINFTTSMRNG..SFCALILHFRPDLIDYKSLNPQDIKENNKAYAGCPASIG 360
Qy 361 ISRLSESDWVLAIIPKLTVMYLYQIRAFSGQELNVQIENSSSTYKVENYEDT 420
Db 361 ISRLSESDWVLAIIPDLITMYLYQIRAFSGQELNVQIENSSSTYKVENYEDT 420
Qy 421 NSSVDQEKFAELSDLKREPELQPISGAVDFLSQDVSFVNDGSGVSESEHQTDPDDL 480
Db 421 NSSVDQEKFAELSDLKREPELQPISGAVDFLSQDVSFVNDGSGVSESEHQTDPDDL 480
Qy 481 SPSTASPYCRRTKSDTEPQKSGSSGRTSGSDDPGICSNSTDSQAQVLLGKRLIKAE 540
Db 481 SPSTASPYCRRTKSDTEPQKSGSSGRTSGSDDPGICSNSTDSQAQVLLGKRLIKAE 540
Qy 541 ELSPDLYSDKKKMDSPICETDQKQLQTLDISNLEKELEKRSRLSECSDESPKX 600
Db 541 ELSPDLYSDKKKMDSPICETDQKQLQTLDISNLEKELEKRSRLSECSDESPKX 600
Qy 601 TSLSPSTKLGYSYRDLDAKKKQASLRQTESDPDADRTTLNHADHSKIVQHRLSROE 660
Db 601 TSLSPSTKLGYSYRDLDAKKKQASLRQTESDPDADRTTLNHADHSKIVQHRLSROE 660
Qy 661 ELKERAVLLEQARDAALKAGNKNTNTATPCNRQLSDQDEERRRLERARQILAE 720
Db 661 ELKERAVLLEQARDAALKAGNKNTNTATPCNRQLSDQDEERRRLERARQILAE 720
Qy 721 ARSGVKNSELPSTYGEAAEKLEKSKASGENDNIEIDTNEIEPEGVVGSGDELTYLEN 780
Db 721 ARSGVKNSELPSTYGEAAEKLEKSKASGENDNIEIDTNEIEPEGVVGSGDELTYLEN 780
Qy 781 DLDPTEONSKLVLDLKKLEVPQVANSPPSSAQKAVTESSEODMSGTEDLTERLQK 840
Db 781 DLDPTEONSKLVLDLKKLEVPQVANSPPSSAQKAVTESSEODMSGTEDLTERLQK 840
Qy 841 TTERFRNPVVSFQSTYRKTQLQSFQYIENRPMKQORSIQEDTKKGNBEKAAITETOR 900
Db 841 TTERFRNPVVSFQSTYRKTQLQSFQYIENRPMKQORSIQEDTKKGNBEKAAITETOR 900
Qy 901 KPSPDEVLNKGFKQTSQYVGEELALENBOKQIDTRALVKRLRYLMDTGRNTEBBEAM 960
Db 901 KPSPDEVLNKGFKQTSQYVGEELALENBOKQIDTRALVKRLRYLMDTGRNTEBBEAM 960
Qy 961 MOEFMLVNNKGNALIRRMNQLSLEKEHDLERYELNLRELRLAMLAIEDMQTEAQKRE 1020
Db 961 MOEFMLVNNKGNALIRRMNQLSLEKEHDLERYELNLRELRLAMLAIEDMQTEAQKRE 1020
Qy 1021 QLLDELVALVNRKDALVRDLDAQEKQAEEDDEHLERTLEQNGKMAKKEKCVLQ 1076
Db 1021 QLLDELVALVNRKDALVRDLDAQEKQAEEDDEHLERTLEQNGKMAKKEKCVLQ 1076
```

```
RESULT 2
US-09-538-092-1136
; Sequence 1136, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Iolc
; APPLICANT: Menzies, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538, 092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurateSeqFormatter Version 0.9
; SEQ ID NO 1136
; LENGTH: 3433
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
FEATURE:
NAME/KEY: misc feature
LOCATION: (0) - (0)
OTHER INFORMATION: Polypeptide Accession Number P46939
US-09-538-092-1136

Query Match 5.9%; Score 326.5; DB 2; Length 3433;
Best Local Similarity 19.8%; Pred. No. 2.6e-12;
Matches 197; Conservative 159; Mismatches 336; Indels 303; Gaps 37;

Qy 286 RKNVASQSLVWCSEVTKNRYGKITYFTTSMRNGSFCALILHFRPDLIDYKSLNPQDI 345
Db 148 QQTNSEKILLSWRQTRPYSOVAVNLFTTSMRNGSFCALILHFRPDLIDYKSLNPQDI 207
Qy 346 KENNKAYDGPAS-IGISRLSESDWVLAIIPKLTVMYLYQI- 388
Db 208 IERLEHAFSAQTYLGLKLDPEB-VAVLPPDKSLIMYLTSLFEVLPQQVITDAIREV 266
Qy 389 ----RAHPSGOELNVQIENSSSTYKVENYED--TNSVDQEKFAELSDLKREPR 441
Db 267 ETLPRKXKCEBEALINIGTAPBEBHESPRAFPTSTVTEVMDLDSYQALAEV----- 321
Qy 442 LQOPISGAVDFLSQDVSFVNDGSGVSESEHQTDPDDLSPSTAS- 488
Db 322 LTMLSAEDTFQEDD--ISDD-VBEVKKQFATHEAFMMLTFAHQSVSGLQAGNQLI 377
Qy 489 CRTKSDTEPQKQ-----QSGRTSGSDDPGICSNSTDSQAQVL----- 528
Db 378 TQGLSDSEEFIEQWTLNLNARWALRVESMQRSLHD-VLMELQKKQLQQLSAMLT 435
Qy 529 LGKRLIKAEITLSD-----LYVSDKKKMDSP 556
Db 436 LTESRIQKMETCPLDLDVKSLOKLEBHKSLQSDLEABQKANSITMIVIVDENGESA 495
Qy 557 PFI-----CESTDQ--KIQTLDISNLEKELEKRSRLSECSDESPKX 599
Db 496 TALLDQLOKLGSRMTAVCWTEERWNRLOEINI--LMQELLE--QCLLKAWLTEX 548
Qy 600 KTSLSPTSKLGYSYRDLDAKKKQASLRQTESDPDADRTTLNHADHSKIVQ- 652
Db 549 BEALNKVQTSNFDQKELSVVRLLAIKE--DMWKRTQL--DQLSIEGQDVGQLD 602
Qy 653 ----HRLSROBELKER-----ARVLEQAR-RDAA 678
Db 603 NSKASKKINSDESELQRMWSLVQRLEDSNGYQAVAKLGMQIIPQKDLLETFRVAREQA 662
Qy 679 LKAGNKNTNTATPCNRQLSDQDEERRRLERARQILAEARSGVKNSELPSTYGE 735
Db 663 ITTKSGQELPPPPPKKQI--HVDIEAKKKFPAISABELNWLKMTTALQTEIKEYMK 720
Qy 736 MA-ABKLEKSKASGENDNIEIDTNEIEPEGVVGSGDELTYLENDLDPTEONSKLVLD 794
Db 721 MDPTSEKKQLKA-----LEKQREIRLP-----ADELNTQGLVEMQKESGLPTE 767
Qy 795 KLKGLLEVPQVANSPPSSAQKAVTESSEODMSGTEDLTERLQKTERFRNPVVSFQ 854
Db 768 EIKGVLE-----KVSSEKNGVSGHLEDLER-----IQQED 799
Qy 855 STYKQLOQSFQYIENRPMKQORSIQEDTK-----KGNBEK 892
Db 800 INAVFQDLDELKVIKITKEWVGHGTSISBSRQSLPSLKDSCORELTNLGLHPKLEMAR 859
Qy 893 AATTEQRKSEDEVLNKG-----KDTQYVVEBLAALNEBOKQIDTRAA 938
Db 860 ASCGALMSQSPADPFQRGDFLGRYQAVQAEVDRQCH-----LEVELGQPGHAY 912
Qy 939 LVEKRLRYLMDTGRNTE-----BEEANQGEFMLVNNKGNALIRRM 978
Db 913 L--ETLKLKQDVANDSENKQVSLNVLNDAKLYEKLQEKTKLDE--ILENGRPAHLKLA 968
Qy 979 NQLSLEK--EHLDEERYELLNLRELRLAMLAIEDMQTEAQKREQLLDELV----- 1028
Db 969 EETKALEKNVHPVEKLYQGFDDVQG-----KMNKLKVLVSKDLILLEBIALTLEAFEA 1023
```

Qy 1029 --ALVNRKDALVRLDAQEKQAEEDHEHLETLQ 1061  
Db 1024 DSTVLEKMDGVKDFLPMKQAAQSPDAGLQRLDQ 1058

RESULT 3  
US-09-949-016-10918  
; Sequence 10918, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10918  
; LENGTH: 2141  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-10918

Query Match 5.6%; Score 307.5; DB 2; Length 2141;  
Best Local Similarity 20.0%; Pred. No. 2.4e-11;  
Matches 196; Conservative 165; Mismatches 318; Indels 303; Gaps 45;

Qy 285 GRKP-NASOSLLVWCKEYTKYRGVKTINFTTSMNGLSFCALIHFRPDLIDYKSLNPQ 343  
Db 173 GRETSADALLMCMQMTAGPHNVNTFTSSWKGALFNLHKKRPDLIDPKLMD 232

Qy 344 DIKENNKAYD-GFASIGISRLLEPSDMVLAIPKLTVMVTVLYOIRAHFSGQELNVQI 402  
Db 223 NARHNLBAFVNAERKOLGIIPLDPEP-VFTENPDEKIIIVVAVFYHFS--KKKVLAV 289

Qy 403 EENSCKTVYKGVNVEDTNNSSVDQKFFAASDL----- 436  
Db 290 E--GKRVGVYIDHAIETKMI--EKYGLASDLTWTIEQITVLNSKKFANSLTGVQOQ 344

Qy 437 -----KRPELQOPISGAVDFLSQDDSVFVNDGVSSESHOTPDHLSPTAS 486  
Db 345 LQASTVATVEKPRFQF--KGNLEVL-----LFTIGSRMANNOKVTPHD--GKLV 394

Qy 487 PYCRTTSDTEPQSQSSGRTSGSDPDGICSTNDSTOAQVILGKRLKAKYTE-LSDL 545  
Db 395 DINRAMESLERAEYRR-----LALNELIRQEKLEQLARR 430

Qy 546 YVSDKKGMSPFICE-----END-----ROKQOT--DI 573  
Db 431 F--DKAAMRETWSENOLVAQDNFGYDLAANBAKKGHAILEDTAAVERVALLEDL 488

Qy 574 GSNLEKELIENSRLSECRSDPSPKITSPTSGLGYSGRDLDLAKKHAISLQTESD 633  
Db 489 AQELEKENVHQKRTAKKNILRL-----WSYIQ--ELQSRQRETTIAL 534

Qy 634 PDADRTTLNADHASKIYQHRU-----LSROESLKRRAVLEQARRDALAKGNKN 686  
Db 535 QKLPQDMHLSIDWMDERVAHLISAFRGHLEVEDLQKHILM-----EADTAIQ--GDKVK 589

Qy 687 TMTATPF-----CNROL-----SPQODE-----RRQLEBAROL----- 717  
Db 590 AITTAITLFTGKGYQPCDPQVITDRISHLEQCEBELSNMAAGRAQL--EOKSLMKFTW 648

Qy 718 -IAEARSQVKNSELPSYGEAAEK-----LKERSKAGDENNIEIDTWEKPIE--GF 767

Db 649 EMDAEASWIKKE-QIYSSLDYGDLSVLLIQKHKAFEDELNGDLNHLQIFQEAHGM 707  
Qy 768 VVGGDDELTNLENLDP--RQNSKYDLKXKLEVPQVANSPPSSAAQKAVTESSQD 825  
Db 708 VA-----RKQPHQIARIKESVQMDQLKDLAFCCKNLDENPFQFGDADD 758

Qy 826 MMSGTEDELTERLQKTEREPNPPVPSKSTYRTQLOSPFQYIENREMKRQSIQBDT 885  
Db 759 LKAWLQD-----AHLISGSDVQDBGATBALGKHQDFLE---ELSEKGVMEHL 806

Qy 886 KKGNEKKAITEYTRKQSEDEVLNKGFKDTSQYVGBELALENROKQIDTPALVEKRL 945  
Db 807 EQ-----QAQGFPEE-----FRD-SPDVTNRLQALRELYQVQAQADLRQRLQ 849

Qy 946 YLMDGTREESBEAMQEWFLVYKKAALIRMMQSLIEKGNHLERYELLNRELRML 1005  
Db 850 EALDLYVFGHTD-CELM--WGEKEXKLAEMEPDTELEEVQNHFFILDDMKITLM 905

Qy 1006 ALED-----WQ--KTEAQRRREQ----- 1022  
Db 906 TQIDGVNLANSVYESGHPREKRVKQYQDHLNTRMQAFQTLVSRREAVDSALVHNVCV 965

Qy 1023 -----LLDE-----LVALVNRDALVND-----LDAQEKAE- 1050  
Db 966 DEERTSKWITDKTKVVESTKDLGRDLAIIAQKLSGLERDVAAIQARVDALERSQOL 1025

Qy 1051 EDEHLERTLBNQKNMAKKEK 1072  
Db 1026 MDSHEQ--KEDIGQRQKHLE 1045

RESULT 4  
US-09-091-501B-10  
; Sequence 10, Application US/09091501B  
; Patent No. 6518413  
; GENERAL INFORMATION:  
; APPLICANT: Tinsley, Jonathan M  
; APPLICANT: Davies, Kay E  
; TITLE OF INVENTION: Utrrophin gene expression  
; FILE REFERENCE: 620-42  
; CURRENT APPLICATION NUMBER: US/09/091,501B  
; CURRENT FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: PCT/GB96/03156  
; PRIOR FILING DATE: 1996-12-19  
; PRIOR APPLICATION NUMBER: GB 9525962.8  
; PRIOR FILING DATE: 1995-12-19  
; PRIOR APPLICATION NUMBER: GB 9615797.9  
; PRIOR FILING DATE: 1996-07-26  
; PRIOR APPLICATION NUMBER: GB 9622174.2  
; PRIOR FILING DATE: 1996-10-24  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 3433  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (239) .. (250)  
; OTHER INFORMATION: Description of Artificial Sequence: Full length  
; OTHER INFORMATION: utrophin construct; Xaa = unknown  
US-09-091-501B-10

Query Match 5.5%; Score 301.5; DB 2; Length 3433;  
Best Local Similarity 19.6%; Pred. No. 1.1e-10;  
Matches 190; Conservative 165; Mismatches 363; Indels 251; Gaps 36;

Qy 286 RKPASOSLLVWCKEYTKYRGVKTINFTTSMNGLSFCALIHFRPDLIDYKSLNPQDI 345  
Db 148 QQNTSRKILSWRQITRYSQVAVNLFTTSMNGLSFALFNAVLRHKKPDLFSMDRVVTKSP 207

Qy 346 KENNKAYD-GFASIGISRLLEPSDMV-----LTAIPDKLT-----VM 382

```

Db      208 IERLEHAFSKAHYIIEKILDPEDVAHLPKXXXXXXXXXXVLEFQVTDIAIREVE 267
Qy      383 TYLQIRAHFSGOELNV---VQIESSSKSTYKVGNETDTNNSVDQKFAEISDLKRE 439
Db      268 TLPRKTKKECEEEHIIQSAVLAEQSP-----RAEPSTVTEVMDMLDSYQIALE 319
Qy      440 PELQOPIGAVDFLSQDVSFVNDGVSSESEHQTPDDHLSPTASPCRTKSDTEPQ 499
Db      320 EVLFWLSAEDTFOEQDD---ISDD-VREVEQFATHETFWMBLTA-----H 362
Qy      500 KSOQSSGRTGSD--DEIGCNTSTQAO---VLGKGLLKATLE---LSDIYSD 549
Db      363 QSSVSVLQAGNQMLTQGLSEEEFEOEQMTLNNRWEALRYVESMERQRLDAMEL 422
Qy      550 KKK---DMSPPFICEETDEQKLTLDIGSNLE--KEKLENSRSECSDSPPIKTSLS 604
Db      423 QKKOLOLSSWALTTEERIQMESPPGLDDLPISQKLIQEKISQNDLEAQ-VKANSLS 481
Qy      605 -----PTSKLGYSYSHDLDAKKKHAHLQ----- 629
Db      482 HMVYIVDENSESATLLEDQQLGGERMTAVCRWTEERNRLOEISILWELLEQCL 541
Qy      630 ----TESDPADRT-TLNADHSSKTYQHLRSQBELKEARVLEQARDALKAQNK 684
Db      542 EAMLTEKEBALNKYQTSNFKQKEISVSRRLALIKEDMEKRTQLDLS-----EIGOD 596
Qy      685 HNTWTATPFCNRQSLDDQDEERR-----QLRRARQOLIAERAGVMSELPSYGEAAE 739
Db      597 VGQLSPKSKSKKNSSEELTORWDSLQRLSDSNQVTAVALKLSQLPQKDLLETV 656
Qy      740 KLERKSKAS-----GDENDNIEIDTNEEIPGFFVVGGD-----E 774
Db      657 HVRKKGAVKPKQELPPPLGPKKQIHVDI--EAKKFFDAISALNMLIKMTAIGTTE 714
Qy      775 LTNLENDLTPBQNSKLVDLKLLKLEVOQVANSPESSAAQKAVTESSEODMSGTEDLR 834
Db      715 IKEXYMKQDTSEMKKKIKALE-KEQREIRPR-ADELAQTCQILVEQ-----MGKEGLP 765
Qy      835 TER-----LOKTERFRN-----PVVESKQSTYRKQLOLQSFQYIENRPEMKQRS 880
Db      766 TERIKNVLEKVSSEWKNVSOHLDELKRIQLOEINNVFYQDLDELEVITKKEEWMVHTS 825
Qy      881 IOEDTK-----KGNREKAITETQKPSDEVLNKG----- 912
Db      826 ISSBSRQSLPSLKQSCQBELTNLGLHPKLEMARASQALSQSDAPVQGRGDSPLGR 885
Qy      913 -----KTSQYVVBELALENQKQIDTRALVETRLRYLMDTGRNTE----- 955
Db      886 YQAVQEAVEDROQH-----LENELKQOPGHAYL--ETIKTLQDLANDSENKAQVSLNV 936
Qy      956 -----BEEAMWQEWFMVLVKNKQNALIRRNQSLILEK--EHLERRYELANREL 1002
Db      937 LNDLAKYKELQEKKTIDE--ILENQKPALHKLAEETKALEKNVHPVEKLYKQEPDVQ 994
Qy      1003 AMLAIEDMQKTEAQKREQLLDELV-----ALVNRDALVLDLAQEKQAEED 1052
Db      995 G-----KMNGLKVLVSDFHLHLEBIALTLRAFEADSVIEKMDGVDFLMKQQAAGDD 1049
Qy      1053 EHLERTLEQ 1061
Db      1050 AGLQROLQ 1058

```

```

RESULT 5
US-09-513-999C-7173
; Sequence 7173, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclerc, A. Y.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961

```

```

; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 7173
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 16
; OTHER INFORMATION: Xaa=Ile or Lys or Arg or Thr
; FEATURE:
; NAME/KEY: UNSTRE
; LOCATION: 17
; OTHER INFORMATION: Xaa=Asp or Glu
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 39
; OTHER INFORMATION: Xaa=Asp or Glu
; US-09-513-999C-7173

Query Match          5.4%; Score 296; DB 2; Length 61;
Best Local Similarity 95.1%; Pred. No. 1,1e-12;
Matches 58; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      727 MSELPSYGEAAAEKLEKERSKASGDENDNIEIDTNEEIPGFFVVGGBDELTNLENDLTPTE 786
Db      1 MSELPSYGEAAAEKLEKERSKASGDENDNIEIDTNEEIPGFFVVGGBDELTNLENDLTPTE 60
Qy      787 Q 787
Db      61 Q 61

RESULT 6
US-08-841-349-4
; Sequence 4, Application US/08841349B
; Patent No. 5955594
; GENERAL INFORMATION:
; APPLICANT: MISHRA, LOPA
; TITLE OF INVENTION: GENES CODING PROTEINS FOR EARLY LIVER DEVELOPMENT. . .
; FILE REFERENCE: XX/P04/70US0
; CURRENT APPLICATION NUMBER: US/08/841,349B
; CURRENT FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2154
; TYPE: PRT
; ORGANISM: Mus musculus
; US-08-841-349-4

Query Match          5.3%; Score 291.5; DB 1; Length 2154;
Best Local Similarity 20.3%; Pred. No. 2,8e-10;
Matches 194; Conservative 146; Mismatches 326; Indels 291; Gaps 38;

Qy      287 KPNASQSLVWCKEVTKNTKRGVKTNPFTTSMRGLSFCAILHFRPDLIDYKSLNPDQIK 346
Db      159 KKSAXDALLMCMQMTKAGYPNVVIHNFITSMKDGMAFNLLIHGRPDLIDPDKLSKNNAH 218
Qy      347 ENNKAYD-GFASIGISRLLEPSDVLALIPDKLTVMTYLYQIRAHFSGOELNVQ---- 401
Db      219 YNLQNAFNLAEOHIGLTKLDPED-LSVDHPDEKSIITVVTYYHYFSKKKALAVEGKRI 277
Qy      402 --IEENSSKSTYKVGNET-----DTNNSV 424
Db      278 GKVLDAINAEIKETKIEYETLASDLLEWIBQTIILNNRFRANSVLGVQOOLAFTYRTV 337
Qy      425 DQEKFAEISDLK-----REPLOOPIGAVDFLSQDVSFVND 464

```

```

Db      338 EKPPKTEKGNLEVLPAIOSKRRANQKVMPEBKGLSIDINKAMERLEK----- 388
Qy      465 GVGSESEHQTPDDHLSPTASPYCRKTSPTPEPOKSOOSSR-----TSSDDPGICS 518
Db      389 --AHEHEBELALNNELIROEKLSQLARFRKAMERETWLSNQRLVSQDNFGFDLPAYEA 446
Qy      519 NTDSQA-----QVLLGKRLKAKET-----LEL----- 542
Db      447 ATKKEALETDIAAVERVQAVVAARELEANNHDIKRTARKONVIRLMEYILLELRA 506
Qy      543 -----SDLYSDKKKMSPPICE-----TDBOKLQTLIDIGS 575
Db      507 RRORLEMLGLQKIFQEMLYIMDMWDEMKVLLISQDYGHLLGVEDLLQKALVEADTAI 566
Qy      576 NLEKKELENSRLSECRSDE--SPYKTSU--SPTSKLGYSRDLDAKKHNASIRQT-- 630
Db      567 QABRVGVNAAQKAPATGEGYKPCDPVIRDRVAHMEFCYOELCOLAERARLEBSRR 626
Qy      631 -----ESDPDADRTTLNHADHSKIVQ--HRLLSR-----OEBLEKARVLLBOAR 674
Db      627 LMKPFWEAABEGMIRREKELISSDDYKGLTSVWELSKHAFEDENSGSG--HFBOAI 685
Qy      675 RDAALKAQNKNTATPFCNRQLSDQDEERRQLREARQLIAEASGVWMSLPSYG 734
Db      666 KEGEDMIAEEN-----FGSEKIRERIIYIRQMANL--EQLSAIRKRLER-- 729
Qy      735 EMAAEKLKERSKASDENNDIEIDTNEBIPGFFVGGD-----ELTNL 778
Db      720 ---ASLHQFOADADDIDAMWLDILKIVSNDV--GHDEVSTQSLVKKHDAVEBITNC 783
Qy      779 ENDDLT--PEONSKYVDLKLKLEVOPOV--ANSPPSAQKAVTESSEODMSGTEDLATE 836
Db      784 RPTIDTLHEQASAL-----POAHASPVVKGRLAIGERCKEMALTY--LKKQ 830
Qy      837 RLQKTERFRPVPVFSKOSTVAKTQLOFSQYIENRPEMKQRSIOEDTKKNEKAIT 896
Db      831 ALQDTLALYK--MFS-----EADACELMIDKEQMLNNMQIP----- 866
Qy      897 ETORRPSDEVLANKFKQTSQYVGBLALENEQKQIDTPRALVEKRLRYLMDTGRNTEE 956
Db      867 ---KLDBLEVIQHRFE-----SLEPEMNQASRAVAVNOIARQLMHNHGPSEK 911
Qy      957 EEAAMQO-----WF-----MLVKNKALIRMNQSLSEKHNLEBRYELNRELAMLA 1007
Db      912 EIRAOQDKLNTRWQSFRELVDKRDAL--LSALSI-----QNYHLECNETKSCIR- 959
Qy      1008 EDWQTEAQKREBOLLND--ELVALVNRDALVRLDAOE-----KOAE--EDEH 1054
Db      960 ---EKTIVIESQDGLNDLAGWMLQCKLTGMERDLVAIEAKLSDLQKEAKLESEH 1013

```

```

RESULT 7
US-09-431-184A-4.
; Sequence 4, Application US/09431184A
; Patent No. 6642362
; GENERAL INFORMATION:
; APPLICANT: MISHA, LOEA
; TITLE OF INVENTION: GENES CODING FOR EARLY LIVER DEVELOPMENT...
; FILE REFERENCE: P04470US1/BAS
; CURRENT APPLICATION NUMBER: US/09/431,184A
; CURRENT FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US98/08656
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: US 08/841,349
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 2154
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-431-184A-4

```

```

Query Match      5.3%; Score 291.5; DB 2; Length 2154;
Beet Local Similarity 20.3%; Pred. No. 2.8e-10;
Matches 194; Conservative 146; Mismatches 326; Indels 291; Gaps 38;

Qy      287 KENASQSLVWCKEVTKNYRQVKTITNFTTSWRNGLSFPCAILHFRPDLIDYKSLNPQDIK 346
Db      159 KKSADALLMWCQMKTAGYPRVNIHNFTTSWRDGAFAALHKKRFDLIDPDKLKSNAH 218
Qy      347 ENNKAYD--GPAISGISHLSEPSDMVLATIDKLTWMTYLYQIRAHFSGELANVQ---- 401
Db      219 YNLQVAFMLABOHGLTGLDLPED--ISVDHDEKSIITVYVYHFFSKMALANEGRI 277
Qy      402 --IENSSKSYKYGNVET-----DTNSGV 424
Db      278 QKVDNALIEKMTIKETTLASDLLEWIBQTIILNKRKANSVLGVQOOLQAFNYRTV 337
Qy      425 DOEKFYAEISDLK-----REPLQCFISGAVDFLSODDSVFVND 464
Db      338 EKPPKTEKGNLEVLPAIOSKRRANQKVMPEBKGLSIDINKAMERLEK----- 388
Qy      465 GVGSESEHQTPDDHLSPTASPYCRKTSPTPEPOKSOOSSR-----TSSDDPGICS 518
Db      389 --AHEHEBELALNNELIROEKLSQLARFRKAMERETWLSNQRLVSQDNFGFDLPAYEA 446
Qy      519 NTDSQA-----QVLLGKRLKAKET-----LEL----- 542
Db      447 ATKKEALETDIAAVERVQAVVAARELEANNHDIKRTARKONVIRLMEYILLELRA 506
Qy      543 -----SDLYSDKKKMSPPICE-----TDBOKLQTLIDIGS 575
Db      507 RRORLEMLGLQKIFQEMLYIMDMWDEMKVLLISQDYGHLLGVEDLLQKALVEADTAI 566
Qy      576 NLEKKELENSRLSECRSDE--SPYKTSU--SPTSKLGYSRDLDAKKHNASIRQT-- 630
Db      567 QABRVGVNAAQKAPATGEGYKPCDPVIRDRVAHMEFCYOELCOLAERARLEBSRR 626
Qy      631 -----ESDPDADRTTLNHADHSKIVQ--HRLLSR-----OEBLEKARVLLBOAR 674
Db      627 LMKPFWEAABEGMIRREKELISSDDYKGLTSVWELSKHAFEDENSGSG--HFBOAI 685
Qy      675 RDAALKAQNKNTATPFCNRQLSDQDEERRQLREARQLIAEASGVWMSLPSYG 734
Db      666 KEGEDMIAEEN-----FGSEKIRERIIYIRQMANL--EQLSAIRKRLER-- 729
Qy      735 EMAAEKLKERSKASDENNDIEIDTNEBIPGFFVGGD-----ELTNL 778
Db      720 ---ASLHQFOADADDIDAMWLDILKIVSNDV--GHDEVSTQSLVKKHDAVEBITNC 783
Qy      779 ENDDLT--PEONSKYVDLKLKLEVOPOV--ANSPPSAQKAVTESSEODMSGTEDLATE 836
Db      784 RPTIDTLHEQASAL-----POAHASPVVKGRLAIGERCKEMALTY--LKKQ 830
Qy      837 RLQKTERFRPVPVFSKOSTVAKTQLOFSQYIENRPEMKQRSIOEDTKKNEKAIT 896
Db      831 ALQDTLALYK--MFS-----EADACELMIDKEQMLNNMQIP----- 866
Qy      897 ETORRPSDEVLANKFKQTSQYVGBLALENEQKQIDTPRALVEKRLRYLMDTGRNTEE 956
Db      867 ---KLDBLEVIQHRFE-----SLEPEMNQASRAVAVNOIARQLMHNHGPSEK 911
Qy      957 EEAAMQO-----WF-----MLVKNKALIRMNQSLSEKHNLEBRYELNRELAMLA 1007
Db      912 EIRAOQDKLNTRWQSFRELVDKRDAL--LSALSI-----QNYHLECNETKSCIR- 959
Qy      1008 EDWQTEAQKREBOLLND--ELVALVNRDALVRLDAOE-----KOAE--EDEH 1054
Db      960 ---EKTIVIESQDGLNDLAGWMLQCKLTGMERDLVAIEAKLSDLQKEAKLESEH 1013

```

```

RESULT 8
US-09-538-092-1243
; Sequence 1243, Application US/09538092

```

```
Patent No. 6753314
GENERAL INFORMATION:
APPLICANT: Glot, Loic
APPLICANT: Mannefeld, Traci A.
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
FILE REFERENCE: 15966-542
CURRENT APPLICATION NUMBER: US/09/538, 092
CURRENT FILING DATE: 2000-03-29
PRIORITY APPLICATION NUMBER: 60/127,352
PRIORITY FILING DATE: 1999-04-01
PRIORITY APPLICATION NUMBER: 60/178,965
PRIORITY FILING DATE: 2000-02-01
NUMBER OF SEQ ID NOS: 1387
SOFTWARE: CurationFormat Version 0.9
SEQ ID NO 1243
LENGTH: 2364
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (0)..(0)
OTHER INFORMATION: Polypeptide Accession Number Q01082
US-09-538-092-1243
```

```
Query Match 5.3%; Score 289.5; DB 2; Length 2364;
Best Local Similarity 19.3%; Pred. No. 4.2e-10;
Matches 208; Conservative 171; Mismatches 369; Indels 331; Gaps 43;
```

```
QY 287 KPNASGSLVWCCKEYTKNYRGVKTITFTTSWRNGLSFCAIIHFRPDLIDYKSLNPQDIK 346
D 172 KKSADLMLLMCKMTAGYVNVNHNFTTSWRDGAFAALVHKRPDLIDDKLKSNV 231
QY 347 ENNKAYD-GFASIGISRLSPSDMVLAI PDKLTWVTYLYQIRAHFSGELAVVQ---- 401
D 232 YNLQNAFRLAEQHLGLTGLDLPED-ISVDHDEKSIITVYTYTHYFSKMAALAVEGKRI 290
QY 402 --IEENSSKSTYKVNTEY-----DTNSSV 424
D 291 GKVLDMALIEKMKIEKESLASDLEWIEQTIIILNNRKFANSLVGVOQLQAFNTYTRTV 350
QY 425 DQKFFYALSDLK-----REPEIQPISGAVDFLSDDSVFVND 464
D 351 EKPRKFTKGNLBYLFTIGSKRANNOKVYMPREBKILSIDINXAMERLEK----- 401
QY 465 GVGSSESEHQTPDHLSPSTASPYCRRTKSDTEPQKSGSSGR-----TSGSDPGICS 518
D 402 --AEHEERLALNELIROEKLQELARRDRKARKEITLSENQKLVSGDNFGFPLAVEA 459
QY 519 NTDSQA-----QVLGKGLLKAET-----LEL--- 542
D 460 ATKGHAIETDIAVEERVAQVAVARELEAENYHDIKRITARKONVIRLWEYLLLELRA 519
QY 543 -----SDLYISDKKXMSPPFICEB-----TDEQKLTLDIGS 575
D 520 RROELEMLGLQKIFOEKLYTMDWDEKVLVLSQDYKHLGVEDLLQKTTIVLADIGI 579
QY 576 NLBEKELNSRLSCSDPE--SPIKKTSL-SPTSKGYSRDLDAKKHNASLRQ-- 630
D 580 QAEIVRGVNSAQCFATDGBGYKPCDPOVYIDRYVHNHFECTQELCQLAAERARLESRR 639
QY 631 -----ESDDPDRITLANHADSSKIVQ--HRLLSR-----QEBLEKARVALLBOAR 674
D 640 LMKRFEMAEEREGMIREKEKILSSDDYQKDLTSVWRLLSKRAPEDEMSGSG--HFEQAI 698
QY 675 RDAALKAGNKA---NNTATATPCNRQLS--DQDEERRRLRERA--RQLLAARS--- 723
D 699 KEGEDMAIEHFSGEKIRERLIYIREQWANLEQLSAIRKKRLBEASLLHPOQADADIDA 758
QY 724 -----GVKMSLELSYGMMAEKRSKASGENDNIE--IDPNEB-----IPB----- 765
D 759 WMLDILKIVSSSDVGHDEYTSQSLVKGHKQVABEAIANTPTLTLHBOASALPOEHAESP 818
QY 766 ---GFVVGAGD-----ELTNLE-----NDLDTPT 785
```

```
D 819 DVGRGLGIERRYKVAELFRLRKQALQDTLALYKMFSEADACELWIDEKQWLNMMQIP 878
QY 786 EQNSKLVLDL--KLLLEVOQVANSPPSAQKRYTSSSEQDMKSGED-----LRTERL 838
D 879 E---KLEDLVTOHFRFSLPEBNNOQASRYA--VVOIAQOLMHSIGPSEKEIKRAQDDKL 933
QY 839 OKTERFRNPVVFSGDSTVKTOLQSPS-QYIENRPMKRORSIOEDTKKGNBEKAAITE 897
D 934 NTRWSQFRELVDKQDLALSLSTIQNYHLECNTEKSIIRKTKVIESTODLGNDLAGYMA 993
QY 898 TORK-----PSEDEVLNKFPDTSQYVVGELALENROQIDT----- 935
D 994 LQRLUTGWERDLVAIEAKSLDLQKEAKELSEHPDQALISRLABISDVWEEMKTLTKN 1053
QY 936 -----RAALVERKRLTMTGRTTEBEEMQWEPVLNKKALIRRNQSLER--- 986
D 1054 REASLGASKLQQLRDLDD-----FQSWLSRTQYALASEDMPVTLTEAEKLLT 1102
QY 987 EHLDERREYELNRELRAMLAIEDMQK-----TEAQ-----KRRQLT---LDEL 1027
D 1103 QHE-NIGNEIDNYE-----EDYQKRMDEGEMVYQGTDAQYMLRQRLQALDYGNNEL 1154
QY 1028 VALVKKDALVRDLDAQE-----KQAE-----EBDEH-----LERTLEQKKKMAKKE 1071
D 1155 HKWEMENQNLISQSHAYQQLFRLDTQKAEFLANQGEVYLAHTEMPTTLEGAEALIKQED 1213
```

```
RESULT 9
US-09-695-795A-2
Sequence 2, Application US/09695795A
Patent No. 6808893
GENERAL INFORMATION:
APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
APPLICANT: ROTHSTEIN, Jeffrey D.
APPLICANT: JACKSON, Mandy
APPLICANT: LAM, Glen
APPLICANT: LAM, Robert
APPLICANT: ORLOV, Irina
TITLE OF INVENTION: GLUTAMATE TRANSPORTER ASSOCIATED PROTEINS AND METHODS OF USE THERE
FILE REFERENCE: JHU1650-2
CURRENT APPLICATION NUMBER: US/09/695, 795A
CURRENT FILING DATE: 2000-10-23
PRIORITY APPLICATION NUMBER: US 60/161,007
PRIORITY FILING DATE: 1999-10-23
PRIORITY APPLICATION NUMBER: US 60/206,157
PRIORITY FILING DATE: 2000-05-22
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 2388
TYPE: PRT
ORGANISM: Rattus
US-09-695-795A-2

Query Match 5.2%; Score 286; DB 2; Length 2388;
Best Local Similarity 22.0%; Pred. No. 7.3e-10;
Matches 200; Conservative 143; Mismatches 320; Indels 246; Gaps 41;
```

```
QY 287 KPNASGSLVWCCKEYTKNYRGVKTITFTTSWRNGLSFCAIIHFRPDLIDYKSLNPQDIK 346
D 175 KKSADLMLLMCKMTAGYVNVNHNFTTSWRDGLAFNALVHGRPDLIDFSLKCNNAH 234
QY 347 ENNKAYD-GFASIGISRLSPSDMVLAI PDKLTWVTYLYQIRAHFSGELAVVQEEEN 405
D 235 YNLQNAFRLAEKELGLTKLDLPED-VVNDQPDDEKSIITVATYTYHFFS--KMAALAVE-- 289
QY 406 SSKSTYVGNVYETDTSNVDOEKFYALSDLKABPEIQPISGAVDFLSDDSVFVND 465
D 290 -GKFIKIVLDHMAEAEHLV--EKYBSLASL-----LQWIEBTIGTF--NDRQ 332
QY 466 VGSESESEHQTPDHLSPSTASPYCRRTKSDTEPQKSGSSGRTSGSDPGICSNSTDSTA 525
```

```

333 LANS-----LSGVNQLOQSNYSYRT--VEKPP--KPTKGNL 365
QY 526 QVLT-----GKRLKLAETLELDLVYSDKKOMSPFICENTDEOKLTLDIGS 575
Db 366 EVLFTTISKLRANNQKVTTPREGRLISDINKAMRLKA-----EHERBALKT---E 416
QY 576 NLEKKELENSRLSECRSDPESPICKTSLSPTSKLGYSYRDLDAK-----KGNASLQOT 630
Db 417 LIRQEKLE---QLARFDRKAMRETWLSNQRLVQSQNFGLAALAAVBAVRKHEAL-ET 472
QY 631 ESDPDADTTLNADHSHKLYOH-----RLSRQEE-----LKE-----RAVYL--- 670
Db 473 DIVAYSGVQAVDAVAELAELAEHYHDIKRIARQNNVAVLWDFLEWVAARERILLNLE 532
QY 671 -----EQARDAALAKGNKHN- 688
Db 533 LQVVPDILLYMDMAEKGRLOQSDLAKHLAGEVLDLLQLHLVADIAVOAERVAVSA 592
QY 689 TATPCN-----ROLSDQDEER---RQLREARQLIAEARSQVMSG-----LPSYGE 735
Db 593 SALRFCDGKGYRPGCPQVSEVATLLEQSYALCELAATRRARLEBRRLMRPLMEVGE 652
QY 736 MAEKLEKRSASGDENDNIEITNEEIPSGFVVGSDLTNLENDL-----TPQNSK 790
Db 653 AEAWVREQHLLASAEI-----GRDLTGVLRLNKHMTALREMSGR 693
QY 791 LVDLKTLKLEVOQVANSPPSAQKATSESE--QDMSGTEDELTERTLOCTTERFRPV 849
Db 694 LGPKLTLLEOGQOLVABEGHPGANAQSTRAALQAOERLEALAEERQAOLAQ--ASLY 750
QY 850 VFSKDSIVTKATQLOFSQYIENRPEM-----KROSIQEDTKGNEKALITET 898
Db 751 QFOADANDMEMAVDALRLVSS--PEVGHDEPSTQALAQHRLABETIAHPTLDALAE- 808
QY 899 QKQSEDEVLNKGFQDTSQYVGBELALENEOKQIDTRBALVERKRLVIMDTGRNTEEE 958
Db 809 -----QAALRPALSHTE--VOGRVPTLEQHYEEILOABA-----GERARALE 849
QY 959 AMQEWPL-----VNGKMLIRMNQLSLEKHD---LERRYELNRLRANLA 1006
Db 850 AALPYTMTLSRAGACGLWBEKE---QWNLGLALPERLEDEPVVQORETLEPEMNALAA 906
QY 1007 IEDMQTEAQKREQLT-----LDELVAL---VNRKDALVRDLDAQKQK-----BEED 1052
Db 907 -----RIRAVSDIAQQLKASPRGDRITIGTQEQALNQRFQFRSLAGGKALTALSQIN 962
QY 1053 EHLERTLEQ 1061
Db 963 YHLECTETQ 971

RESULT 10
US-09-949-016-6777
; Sequence 6777, Application US/09949016
; Patent No. 681239
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYOMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6777
; LENGTH: 1026
; TYPE: PRT
```

```

; ORGANISM: Human
US-09-949-016-6777
Query Match 5.2%; Score 285; DB 2; Length 1026;
Best Local Similarity 21.2%; Pred. No. 2.7e-10;
Matches 242; Conservative 153; Mismatches 363; Indels 384; Gaps 58;

QY 116 DDP--DAEELNPFQDP-----DSEERTETASPRTEDESFYNNSNYPPEVOTPOLY 165
Db 79 DGRSDSVESGPFPRPSPSAVDEDEKFIASGTT-----YMLDNIELVDTFQTL 130
QY 166 NP-----FD-----EPEA 173
Db 131 EPRASDAKNGQKVNTRKSTDSVPISKSTLSRLSLQASFPDQASSGNPEAVALARDA 190
QY 174 FVTIKDSPQSTKR-KNIRPDMSKYLYADSSKTEBEBLDSNPPEKSTPPNNLV- 230
Db 191 YSTGSSASSTLKTKTKRPPSLKK-----KOTTKKPTETPPVKETQGFDESLVPS 243
QY 231 --NPVQLETERAVYKRAKAP-----PVLSPKT-----QVLN-----E 261
Db 244 GENLASSETKIE-SAKTIGPSPALIEETPLBPAVGPKAACPIDSASAGVPPASGGGRVQ 302
QY 262 NTVSAK--DLSTSPKSPPI-PPVILGRKPNASQSLVMCKEYTKNYRGVKITFTTSM 317
Db 303 NSPPVGKRTPLPLTAPAGAVTTPSDSGQEDSPAKGLSV---RLBFVDSBK-----SSW 354
QY 318 RNLGSPCALIHFPRDLIDYKSLNPQDIKENNKAYNGFASIGISRLIEPSPDMVLAIPD 377
Db 355 -----DNGQENPPPTKTKGKK-----PVAKMPLRPPK--MIKTPE 387
QY 378 KLTVNTYLYQIRAHPSQGLNVVQIENSSKSTYKGVNRYETDNTSSVQGEK--YAEISD 435
Db 388 KLD-NTPASPR--SPABPDIPY---AGTY--TFDDI---KMDPNFNPSSSTSK 433
QY 436 LKREPELQPISGAVFLSDPDSVFVNDSGVSESEHQTPDDHLSPESTASPYGKRTSD 495
Db 434 MQESPKLPQ-----QSYNFPDD--TCBESVDPFTSSKTP 466
QY 496 TEPOKS-----QOSSGRTSGSDPGICSTNDSTQAOVLAKKRLKLAETLESLDYSD 549
Db 467 SSPGKSPASFEIPIPSAMEANGVDGGL--NKPARK-----KKTPLKTDTRVVK--SP 515
QY 550 KKKMSP-----FICERTDEOKL-----QTLQISLVEKEL 582
Db 516 KRSPLSDPPSQDPPTAETPETPVISAVVHATDEKLAVTNQKTCMTVD---LEADKQ 571
QY 583 ENSRLSECRSDPESPICKTSL-SPTSKLGYSYRDLDAKKGASLQRTESDPADRTTL 641
Db 572 DYPOP-----SDLSTFVNETKFSPTBELDTRNSYIETMEKIGSSLQODDAP----- 620
QY 642 NHADHSHKLYOHRLLSQEBELKERARVLLBOARRDALKAGNKNTATATPCNRQLSDQ 701
Db 621 -----KQALYIMFDTSQSPVKSPPVMSSPPT-----CSGS 654
QY 702 QDEERRQLKERA-----ROLIAKRSQVXSELSYSGMAEKLKESKASGDERDN 754
Db 655 SFETELAVTVAANQHPVPRGLAPNOBSHLQVPEKKSQKLELMGL-----GTSEEA 707
QY 755 IEIDTNEIIEPGFVVGDELITNLENDL-----DTPQNSKLVLDLKKGLLEVQ 803
Db 708 IEITA---PEGSASADALLSLRLAHVSLGALDYLEPLDAEKONPPLFAQKQBELEFEA 763
QY 804 PQVANSPPSAQAQVATSSSQDMKSGTEDELTERTLOKT-----TERFRNP--VVSFK- 853
Db 764 IMRIEALKARQIALASRSHQDAKRABAHPTDVASISKTALYSRIGTAVEKFPAGILFQOP 823
QY 854 --DSTVAKTQLOGS-----QYENRPEKRRRSIQEDTKKNEEALATTEQOK 901
Db 824 DLDSALQIAAEIITTKEREVSWMKDYESSRREVMENRKIVABEYK-----TTAQ 873
QY 902 PSEDEVLNKGFQDTSQYVGBELALENEOKQIDTRBALVERKRLVIMDTGRNTEEBEAM 961
```



```

Db      874 MIEDE---QREKSVSHQTVQOL-VLEKQALADLNS--VEKS---LADLFRYKMKVYL 924
Qy      962 QEWFMVLVKNKQALIRRNQ--LSLEKXHDLEERYELL-----NRELAM--LA 1006
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      925 EGF-----RKBEVYAKCAQETLSVKKE--EQRYQALKVAEKLDRANAELIQVNGKA 977
Qy      1007 IEDMQKTEAQKRRBQLLDLDELVALVNRDALVRDLDAQEKQABEDHELTLEQNGKM 1066
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      978 QOEQAHAQASLRKEQL-----RYDALERTLEQNKKEIE--LTKICDELLAKM 1023
Qy      1067 AK 1068
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1024 GK 1025

```

```

RESULT 11
US-09-949-016-10870
; Sequence 10870, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10870
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10870

```

```

Query Match      5.2%; Score 285; DB 2; Length 1034;
Best Local Similarity 21.2%; Pred. No. 2,7e-10;
Matches 242; Conservative 153; Mismatches 363; Indels 384; Gaps 58;

Qy      116 DDP--DAEELNPFQDP-----DSEPTITERASPKTEDSPFNNSYFPKGVQTOYL 165
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      87 DGRSRDSVESGPRPRPSHSVAVDKPIASSGT-----YNLDFNIELVDTQTQL 138
Qy      166 NP-----FD-----EPEA 173
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      139 EPRASDAKQEGKVNTRKSTDSVPSKSTLSRLSLQASDFDQASSGNPEAVALAPDA 198
Qy      174 FVTIKSPSPQSTGR-KNIRVVDMSKITYADSSKTEEBELDESFPYFKTPPPNNIV-- 230
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      199 YSTGSSASASTLRTKTKPRPSLKK-----KQTKKPTETPVKTEQOEPEBESLVP 251
Qy      231 --NPVOELTERRRYKRAKAPAP-----PVLSPKT-----GVLN-----E 261
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      252 GENLASETKTE-SAKTGGPSFALLERTPLBPAGVFKACACPLDSDSAGVVPVPSGGGRVQ 310
Qy      262 NTVSACK--DLSTSPKSPDI--PSVYGRKPNASQSLVWCKEVTKNYRGVKTITFTTWS 317
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      311 NSPPVGRKTLPLTTAPBAGVTPSDSGGDSPAKGLSV--RLKFPYSDK-----SSW 362
Qy      318 RNLGSLFCALIHRRPDLIDYKSLNPDQIKENNKKAYGFGASIGISRLBESDMVLAI 377
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      363 -----DMQOENPPTPKTKIGKK-----PVAKMPLRRRK--MKKTPS 395
Qy      378 KLTVMYTYLVQIRAHFSGQGLNVVQIBNSSKSTYKVGNYEDTNNSVDQKPF--VATLSD 435
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      396 KLP--NTPASRPR--SPAENDIPF-----AKGTY--TFPID--KWDDENFNPFSTSK 441
Qy      436 LKKEPELQPIGSAVDFLSODDSVFVNDSGVGSSESHQTPDDHLSPSTASPYCRRTKSD 495

```

```

Db      442 MQESPRLQ-----QSVPDPD--TCDESVDPFKTSKTP 474
Qy      496 TEPQKS-----QSSSGRTSGDDPGICSNITDQVAVLGGKRLLAARLELSDLVSD 549
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      475 SSPSKSPASFEIPASAMEANGVDGGL--NKPAK-----KTPPLTDTFRRVK--SP 523
Qy      550 KKKDMSPP-----FICETDEOKL-----QTLIGSLVLEKEL 582
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      524 KRGPLSDPPSODPTPATPETPVPVISAHVATBEKLAIVNQKTCMTVD--LEADKQ 579
Qy      583 ENSRSLECRSDPSPPIKTSL-SPTSXLYGSYSRDDLDLAKKQNASLQTESDDPADRTTL 641
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      580 DYPQP-----SDLSTFVNETKFSSTPEBELDYRNSYEIEYMEKIGSLPQDDAP----- 628
Qy      642 MHADHSKIYQHLLSQEELKERARVLEQARDDALTKAGNGNTTATPFCNRQISDQ 701
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      629 -----KKQALYLMEDTQOESPVKSSPVRMSSPTP-----CSGS 662
Qy      702 QDEERRRQLRERA-----ROLIAEARGVKMSLPSYGEAAEKLKERSKASGDENDN 754
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      663 SFETETALVNTAKNQHPVPRGLAPNQESHLOVPKSSQKELAMGL-----GTSEA 715
Qy      755 IEIOTNEEIPGFGVVGGDLYNLNDL-----DTPQNSKLVLDKIKLIEVQ 803
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      716 IEITYA-----PEGSFASADALLSRLAHVSLGALDYLEPLAEKNPPLFAQKLOEELFA 771
Qy      804 PIVANSPSSAAQAVYNESSQDDKSGTEDRTERLOKT-----TERFRNP--VVPK- 853
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      772 IMRIEALKLARQIALSRSHQDAKREAAHPTDVISISTALYSRIGTAEVKEPGLLEQOP 831
Qy      854 --DSTVAKTQLOSFS-----QYENRPMKQORSIOEPTKKGNEBKAATETQK 901
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      832 DLBSAQIARABITITERBVSWMKDYERSRBRVMEKRLVABYEK-----TIAQ 881
Qy      902 PSEDEVLNKGPQDTSQVVGELALENEQOQIDTRALVEKRLRYLMDTGRNTEEBRAM 961
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      882 MIEDE---QREKSVSHQTVQOL-VLEKQALADLNS--VEKS---LADLFRYKMKVYL 932
Qy      962 QEWFMVLVKNKQALIRRNQ--LSLEKXHDLEERYELL-----NRELAM--LA 1006
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      933 EGF-----RKBEVYAKCAQETLSVKKE--EQRYQALKVAEKLDRANAELIQVNGKA 985
Qy      1007 IEDMQKTEAQKRRBQLLDLDELVALVNRDALVRDLDAQEKQABEDHELTLEQNGKM 1066
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      986 QOEQAHAQASLRKEQL-----RYDALERTLEQNKKEIE--LTKICDELLAKM 1031
Qy      1067 AK 1068
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1032 GK 1033

```

```

RESULT 12
US-09-270-767-42273
; Sequence 42273, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 42273
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-42273

Query Match      5.2%; Score 283.5; DB 2; Length 153;
Best Local Similarity 53.1%; Pred. No. 2.6e-11;
Matches 52; Conservative 18; Mismatches 27; Indels 1; Gaps 1;

```



```

Db 511 PNNMMSAMBDLQDMFIVHTIEBI-EG-LISANDQFCKSTLPDADREBBAIAIHKEAQR 568
Qy 799 LLE-----VOPQVANSPPSSAAOKAVTES-----SEQDKSGEDRTA--- 836
Db 569 IASBNHKLKSGSNPYTTVTPOIINSKWEKVOQLPKDHALLEQSQSQSNEHLRRQFAS 628
Qy 837 -----RLQKTERFRNPVFSKSTVRKTQLOQSFQY-----IENPEM-----KROSI 881
Db 629 QANVVGWIGTKYKEI-GRISIENNGTL-EDQLSHLQKQERSIVDYKPNLDLEQCHOLI 686
Qy 882 QE-----DTKGN-----EERKAITETORRPSSEDEVLNKGFKDTQYVVEGLALE 927
Db 687 QEALIFPNKHTNYMEHIRVGMEOQLTTIARTINEVENQILTRDAKISIQEOMQEFPRASF 746
Qy 928 NEQKOIDTRALVKEKRLRYLMDTGRTTEE---EAMQEWMLVYNKKA 973
Db 747 NHFDKQHGALGPEFRACLISLGYDVENDRQGEAEFRIMSLVDPMHS 795

```

## RESULT 15

```

US-09-091-501B-8
; Sequence 8, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathan M
; APPLICANT: Davies, Kay E
; TITLE OR INVENTION: utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091.501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 8
; LENGTH: 2008
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (239) ... (250)
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric; Xaa = unknown
US-09-091-501B-8

```

```

Query Match 4.98; Score 270.5; DB 2; Length 2008;
Best Local Similarity 20.66; Pred. No. 6.1e-09;
Matches 211; Conservative 164; Mismatches 349; Indels 301; Gaps 53;

```

```

Qy 286 RKNPASOSLWCKEYTKNRYGVKITFTSMRNGSLFSCILHNFRLIDYSLNFDI 345
Db 148 QQTNSKILISWVRQTRPFSQVAVNLFTTSWIDGLAFNAVLRHKRDLFSMDRVVMSP 207
Qy 346 KENNKAAYD-GFASIGISRLLEPSDNV-----LAIAPDKLT-----VM 382
Db 208 IERLEHAFSAHTYUIGIEKLDPEDEVVHL.PXXXXXXXVYEVLPQVTTIDALREVE 267
Qy 383 TTYIQIRAHSGOELNV---VOIEENSSKSTYKGVNETDTNNSVDOEKFYAELSLKRE 439
Db 268 TLPRKTKKECEBEIHIQSAVLAEBSQSP-----RAETPSTVEVMDLDSYQIALE 319
Qy 440 PELQOPISGAVDFLSODD-----SVFVNDSGVGE----- 468
Db 320 EVLTWLSADTQOEQDDIDDDVBEVKEQPATHETFMELTAHQSSVGLQAGNQLMTQ 379
Qy 469 ---SESEHQTPDDHLSBSTASPYCRATKS-----DTEPOKSGQSSG----- 506

```

```

Db 380 GTLSEEEFEIQQEOMTLINARWEALRVESMEROSRLDALMLEQKQLOQLSSWLTALTEB 439
Qy 507 RTSGSDDPGICSNWTSQAOVNLGKKRL---LKAETLESLD-----Y 546
Db 440 RIQRMESPPLGDDLPSTIQ-KLQEHKSLQNDLEAEQVKNVSLTHMVIVDENSEGSAITAL 498
Qy 547 VSDKQMSPPF--ICEETDEQ--KLQTLDIGNLKEKLENSRSLEC-RSDESPXYKT 601
Db 499 LEDQLKLGRRMAVACWMEBERNNRQIEISI---LWGEELLEQCCLLEAWLTREBRALNV 555
Qy 602 SLSPSTSLGYSYRRDLDAKKKIASLRQTESDPDADRTTYLNHADHSKIYQ----- 652
Db 556 QTS-----NEKDDKEISVSVRRLAIKE---DHEMKQTL---DQLSIQQDVQQLSNP 604
Qy 653 ---HLLSROELKEBRARVLEQAR-----RDAALKAG----- 682
Db 605 KASKKMSDSEELTORWDSLVQRLEDSNQTQAVAGLGSQIPQDLETFVHREKGV 664
Qy 683 NKENWTATPFQCNRLSDQODE-----ERRQLREARQL--IAEASGVMSLPSYGE 735
Db 665 KKKQKQELPPLTKAHHAMQKRSITTELGENTQELRLQKEMVAAKLMINRTL-----E 720
Qy 736 MAAEK---LKERKASGDENDNIEIDTNE--BIPEGVVGGDDELTLNENDLTPRQNS 789
Db 721 MLSCKSLSPERDKIS-ESLRVTNMTWNKICREVP-----TTLKECIEPSSVS 768
Qy 790 -----KLVDLKTKLLEVOQOVANSPPSAA--OKAVTES-----RQDKSGT 830
Db 769 QTRIAHPNVQKVLVSSASADIVQSHRTSEISIPADLDITTELMDVLIDQMLKSN 828
Qy 831 EDL-RTERLOKTERFRNPVFSKSTVRKTQLO---SFSQYENR---PEMK----- 876
Db 829 VVYGVDEBEINKYTSRK---ITGADLEQNHQPDYFTYLAQNLKKAASSDMRTAITEKL 885
Qy 877 -----RQRSIQE---DTKKGNEKKAITETORK-----PSEDEV 908
Db 886 ERYVGNQWDGTQHGVELRQQLLEMIIDSLQWDHRETEBELMKYEALTYILQARDDL 945
Qy 909 NKQPKDTQYVVEGLAALBENQOIPTRALVKEKL-RYLMDSGRTEE-EEMQEWPM 966
Db 946 TKQISD-NQTLLOELGPGDGIYVAFDN---VLQKLEEGSDTRNVKETEYLTQTSW-- 999
Qy 967 LVNKNALIRRMNQSLLEKEHDLERRYBELNRELAMLAIEDM-OKTEA-----QKR 1018
Db 1000 -IMKQSIADRON--ALBAEW--RTVQASRRDLNPL--KMIQEAEFTVVAVLDASH 1049
Qy 1019 REQLLDELVALVKNKDALVRDIDAQEKQAESE-DEH--LERTLEQNKGMAR---KEE 1071
Db 1050 RENVLQDSILA-----RELKQOMODIQAEIIDAHNDFIKSIDGNRQRMVYALGNSSE 1100
Qy 1072 KCVLQ 1076
Db 1101 ATMLQ 1105

```

```

Search completed: May 20, 2006, 17:16:29
Job time : 60 secs

```

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

## OM protein - protein search, using sw model

Run on: May 20, 2006, 17:26:58 ; Search time 186 Seconds  
(without alignments)  
2679.673 Million cell updates/sec

Title: US-09-856-723a-8

Perfect score: 5489  
Sequence: 1 MKQYASMPQTVDVQLKFKP.....RTLEHNGKMAKKEKCVLIQ 1076

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

Published Applications\_MA Main:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/pubppaa/US07\_PUBCOMB.pep:\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/pubppaa/US08\_PUBCOMB.pep:\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/pubppaa/US10A\_PUBCOMB.pep:\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/pubppaa/US10B\_PUBCOMB.pep:\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/pubppaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5486	99.9	1076	4 US-10-275-595A-26	Sequence 26, Appl
2	5173	94.2	1023	5 US-10-450-763-53242	Sequence 5342, A
3	4485	81.7	882	4 US-10-408-765A-1942	Sequence 1942, Ap
4	3151	57.4	669	5 US-10-450-763-53241	Sequence 53241, A
5	1045	19.0	221	5 US-10-450-763-53239	Sequence 53239, A
6	1020.5	18.6	957	6 US-11-097-143-39276	Sequence 39276, A
7	989	18.0	479	5 US-10-450-763-53240	Sequence 53240, A
8	945.5	17.2	1523	4 US-10-473-576-5	Sequence 5, Appl1
9	897	16.3	175	3 US-09-864-761-42978	Sequence 42978, A
10	897	16.3	175	4 US-10-028-386-33821	Sequence 33821, A
11	510.5	9.3	996	6 US-11-097-143-23565	Sequence 23565, A
12	507.5	9.2	1010	4 US-10-359-012-18	Sequence 18, Appl
13	498.5	9.1	863	4 US-10-359-012-2	Sequence 2, Appl1
14	498.5	9.1	863	4 US-10-359-012-14	Sequence 14, Appl
15	498.5	9.1	863	6 US-11-169-041-167	Sequence 167, App
16	461	8.4	91	3 US-09-864-761-34415	Sequence 34415, A
17	461	8.4	91	3 US-09-864-761-46534	Sequence 46534, A
18	441	8.0	1480	4 US-10-359-012-22	Sequence 22, Appl
19	431	7.9	1026	4 US-10-359-012-23	Sequence 23, Appl
20	429.5	7.8	904	4 US-10-359-012-16	Sequence 16, Appl
21	418	7.6	79	3 US-09-864-761-37001	Sequence 37001, A
22	418	7.6	79	3 US-09-864-761-46885	Sequence 46885, A
23	399.5	7.3	1633	4 US-10-359-012-4	Sequence 4, Appl1
24	368.5	6.7	1004	4 US-10-291-172-235	Sequence 235, App
25	368.5	6.7	1004	4 US-10-221-278-235	Sequence 6, Appl
26	368.5	6.7	1965	4 US-10-359-012-6	Sequence 611, App
27	366.5	6.7	1025	4 US-10-291-172-611	

28	366.5	6.7	1025	4 US-10-221-278-611	Sequence 611, App
29	333	6.1	791	4 US-10-170-385-57	Sequence 57, Appl
30	327.5	6.0	1452	5 US-10-995-561-778	Sequence 778, Appl
31	326.5	5.9	3433	4 US-10-408-765A-731	Sequence 731, App
32	326.5	5.9	3433	4 US-10-408-765A-732	Sequence 732, App
33	326.5	5.9	3433	5 US-10-756-149-5681	Sequence 5681, App
34	319	5.8	65	4 US-10-425-115-333910	Sequence 333910,
35	319	5.8	1790	5 US-10-732-923-3331	Sequence 3331, Ap
36	318.5	5.8	1268	5 US-10-732-923-3332	Sequence 3332, Ap
37	318.5	5.8	1790	4 US-10-369-493-1586	Sequence 1586, Ap
38	318.5	5.8	1790	5 US-10-732-923-3330	Sequence 3330, Ap
39	316.5	5.8	2291	6 US-11-097-143-12420	Sequence 12420, A
40	316.5	5.8	4101	4 US-10-369-493-6509	Sequence 6509, Ap
41	316.5	5.8	4101	4 US-10-369-493-6510	Sequence 6510, Ap
42	309	5.6	2326	4 US-10-369-493-6374	Sequence 6374, Ap
43	308.5	5.6	2137	5 US-10-756-149-4806	Sequence 4806, Ap
44	307.5	5.6	2106	4 US-10-408-765A-2093	Sequence 2093, Ap
45	305.5	5.6	915	6 US-11-097-143-8538	Sequence 8538, Ap

## ALIGNMENTS

```
RESULT 1
US-10-275-595A-26
; Sequence 26, Application US/10275595A
; Publication No. US20040078804A1
; GENERAL INFORMATION:
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Dyrng Alma M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: HITLMAN, Jennifer L.
; APPLICANT: AZIMZAI, Valda
; APPLICANT: LAU, Preeeli
; APPLICANT: YAO, Montique G.
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: BATRA, Satjeet
; APPLICANT: KEARNEY, Liam
; APPLICANT: POLICKY, Jennifer L.
; TITLE OF INVENTION: CYTOSKELETON-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0772 USN
; CURRENT APPLICATION NUMBER: US/10/275,595A
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/201,960
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/202,729
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: US 60/209,705
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 60/210,149
; PRIOR FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: US 60/213,215
; PRIOR FILING DATE: 2000-06-21
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PERL Program
; SEQ ID NO 26
; LENGTH: 1076
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 5202390CD1
US-10-275-595A-26
Query Match 99.9%; Score 5486; DB 4; Length 1076;
Best Local Similarity 99.9%; Pred. No. 1.9e-287;
Matches 1075; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
1 MKQYASMPQTVDVQLKFKPVSAAALQFSISCTFLREGKATDSDMOSIASVSKQ 60
|||||
```

Db 1 MKQYASMPWQTVDVKKFKPLSKRVSAALQFSLSCIFLREKATDEDMQSLASIMSKQ 60  
QY 61 ADIGNLDPEEDNEDDENRNVNOEKAKITELINKLNFLEBAEKDLATVNSNPPDDDA 120  
Db 61 ADIGNLDPEEDNEDDENRNVNOEKAKITELINKLNFLEBAEKDLATVNSNPPDDDA 120  
QY 121 AELNPFDPDSEEBITETASPRKTEDSFYNNSYNPFKEVQTPQYLNPFDEBAFVTKDS 180  
Db 121 AELNPFDPDSEEBITETASPRKTEDSFYNNSYNPFKEVQTPQYLNPFDEBAFVTKDS 180  
QY 181 POSTKRNIRPVMSKYLVAADSKTEBEELDESNPFYEPKSTPPNNLVNPFVLETER 240  
Db 181 POSTKRNIRPVMSKYLVAADSKTEBEELDESNPFYEPKSTPPNNLVNPFVLETER 240  
QY 241 RVKRAKAPPVLSKPTGVLENENTVSAGKDLSTSPKSPISPVPVGRKPNASQSLVWCKE 300  
Db 241 RVKRAKAPPVLSKPTGVLENENTVSAGKDLSTSPKSPISPVPVGRKPNASQSLVWCKE 300  
QY 301 VTKXNKGKLTNFTTSWRNGLSFCALLHFRPDLIDYKSLNPDIKENNKKAYDGFASIG 360  
Db 301 VTKXNKGKLTNFTTSWRNGLSFCALLHFRPDLIDYKSLNPDIKENNKKAYDGFASIG 360  
QY 361 ISRLLESMDVLAIPDKLTVMYTLVQIRAHFSGQELNVVQIENSSKSTYKGVNRETDT 420  
Db 361 ISRLLESMDVLAIPDKLTVMYTLVQIRAHFSGQELNVVQIENSSKSTYKGVNRETDT 420  
QY 421 NSVDQEKFYAELSDLKREPELOPISGAVDFLSQDSDVFVNDGVSSESEHQTDPDHL 480  
Db 421 NSVDQEKFYAELSDLKREPELOPISGAVDFLSQDSDVFVNDGVSSESEHQTDPDHL 480  
QY 481 SPSTASPYCARTKSDTEPQKSSQSSGRTSGSDPGICSNSTSTQAVYLLGKKRLKAETL 540  
Db 481 SPSTASPYCARTKSDTEPQKSSQSSGRTSGSDPGICSNSTSTQAVYLLGKKRLKAETL 540  
QY 541 ELSDLYVSDKKKMSPPFICEETDEQKLTLDIGSNLEKELBNSRSLBECRSPBESPIKK 600  
Db 541 ELSDLYVSDKKKMSPPFICEETDEQKLTLDIGSNLEKELBNSRSLBECRSPBESPIKK 600  
QY 601 TSLSPSTKLGYSYRDLDAKKQKASLRQTESPDADRTLNADHSSKTVORHLLSROE 660  
Db 601 TSLSPSTKLGYSYRDLDAKKQKASLRQTESPDADRTLNADHSSKTVORHLLSROE 660  
QY 661 ELKERARVLLQARRDALAKGNKHNNTATPFCNRQLSDQDEERRRLREBARQLIAE 720  
Db 661 ELKERARVLLQARRDALAKGNKHNNTATPFCNRQLSDQDEERRRLREBARQLIAE 720  
QY 721 ARSGVMSKELPSYGEAAAEKLERKSKASGDENDNIEITDNEEIPBEGTVGSGDELTVLEN 780  
Db 721 ARSGVMSKELPSYGEAAAEKLERKSKASGDENDNIEITDNEEIPBEGTVGSGDELTVLEN 780  
QY 781 DLDTPEONSKLVLDLKKGLLEVOPOVANSPSAAOKAVTSSBODMKSGETEDLRTERLOK 840  
Db 781 DLDTPEONSKLVLDLKKGLLEVOPOVANSPSAAOKAVTSSBODMKSGETEDLRTERLOK 840  
QY 841 TTERFRNPVVSFSDSTVRKTLQOSFSQYIENRPEMKRQRSIOBPTYKGNKEKAITETOR 900  
Db 841 TTERFRNPVVSFSDSTVRKTLQOSFSQYIENRPEMKRQRSIOBPTYKGNKEKAITETOR 900  
QY 901 KPSBEDVLNKGFDOTQYVVGELAALENQKQIDTTRALVEKRLRYLMDTGRNTEBEBAM 960  
Db 901 KPSBEDVLNKGFDOTQYVVGELAALENQKQIDTTRALVEKRLRYLMDTGRNTEBEBAM 960  
QY 961 MOBEMFLVNNKQALIRRMQSLLEBKHDEERYELLNRELRLAMLAIBDQKTEAQRRRE 1020  
Db 961 MOBEMFLVNNKQALIRRMQSLLEBKHDEERYELLNRELRLAMLAIBDQKTEAQRRRE 1020  
QY 1021 QLLLDLVALVNRDALVNRDLAQAQKQAEEDHLEKTLLEQNKGNKAKKEKCVLQ 1076  
Db 1021 QLLLDLVALVNRDALVNRDLAQAQKQAEEDHLEKTLLEQNKGNKAKKEKCVLQ 1076

RESULT 2  
US-10-450-763-53242

Sequence 53242, Application US/10450763  
Publication No. US20050196754A1  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc  
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
FILE REFERENCE: 790CIP3/US  
CURRENT APPLICATION NUMBER: US/10/450,763  
CURRENT FILING DATE: 2003-06-11  
PRIOR APPLICATION NUMBER: PCT/US01/08631  
PRIOR FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: 09/540,217  
PRIOR FILING DATE: 2000-03-31  
PRIOR APPLICATION NUMBER: 09/649,167  
PRIOR FILING DATE: 2000-08-23  
NUMBER OF SEQ ID NOS: 60736  
SOFTWARE: Custom  
SEQ ID NO 53242  
LENGTH: 1023  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (254)..(284)  
OTHER INFORMATION: Actinin-type actin-binding domain proteins domain identified  
OTHER INFORMATION: by eMATRIX, accession number BL00019D, p-value=4.200e-19, raw score  
OTHER INFORMATION: of 15.33  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (237)..(342)  
OTHER INFORMATION: Calponin homology (CH) domain identified by Pfam, accession  
OTHER INFORMATION: name CH, E-value=2.4e-17, Pfam score of 71.1  
US-10-450-763-53242

Query Match 94.2%; Score 5173; DB 5; Length 1023;  
Best Local Similarity 99.7%; Pred. No. 1,4e-270;  
Matches 1011; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 63 IGNLDPEDNEDDENRNVNOEKAKITELINKLNFLEBAEKDLATVNSNPPDDDAE 122  
Db 10 IGNLDPEDNEDDENRNVNOEKAKITELINKLNFLEBAEKDLATVNSNPPDDDAE 69  
QY 123 LNPFGDPDSEEBITETASPRKTEDSFYNNSYNPFKEVQTPQYLNPFDEBAFVTKDSPP 182  
Db 70 LNPFGDPDSEEBITETASPRKTEDSFYNNSYNPFKEVQTPQYLNPFDEBAFVTKDSPP 129  
QY 183 QSTKRKNIRPVMSKYLVAADSKTEBEELDESNPFYEPKSTPPNNLVNPFVLETERRY 242  
Db 130 QSTKRKNIRPVMSKYLVAADSKTEBEELDESNPFYEPKSTPPNNLVNPFVLETERRY 189  
QY 243 KRAKAPPVLSKPTGVLENENTVSAGKDLSTSPKSPISPVPVGRKPNASQSLVWCKEYT 302  
Db 190 KRAKAPPVLSKPTGVLENENTVSAGKDLSTSPKSPISPVPVGRKPNASQSLVWCKEYT 249  
QY 303 KNYRGVKTNFTTSWRNGLSFCALLHFRPDLIDYKSLNPDIKENNKKAYDGFASIGIS 362  
Db 250 KNYRGVKTNFTTSWRNGLSFCALLHFRPDLIDYKSLNPDIKENNKKAYDGFASIGIS 309  
QY 363 RLLEPSDMVLAIPDKLTVMYTLVQIRAHFSGQELNVVQIENSSKSTYKGVNRETDTNS 422  
Db 310 RLLEPSDMVLAIPDKLTVMYTLVQIRAHFSGQELNVVQIENSSKSTYKGVNRETDTNS 369  
QY 423 SVDOEKFYAELSDLKREPELOPISGAVDFLSQDSDVFVNDGVSSESEHQTDPDHLSP 482  
Db 370 SVDOEKFYAELSDLKREPELOPISGAVDFLSQDSDVFVNDGVSSESEHQTDPDHLSP 429  
QY 483 STASPYCARTKSDTEPQKSSQSSGRTSGSDPGICSNSTSTQAVYLLGKKRLKAETTEL 542  
Db 430 STASPYCARTKSDTEPQKSSQSSGRTSGSDPGICSNSTSTQAVYLLGKKRLKAETTEL 489  
QY 543 SLDLYVSDKKKMSPPFICEETDEQKLTLDIGSNLEKELBNSRSLBECRSPBESPIKKT 602  
Db 490 SLDLYVSDKKKMSPPFICEETDEQKLTLDIGSNLEKELBNSRSLBECRSPBESPIKKT 549

```

Qy 603 LSPYTSKLGYSYRDLAKKKHASTLRQTESPPDADRTTLNHADSSKIVQHRLLSROBEL 662
Db 550 LSPYTSKLGYSYRDLAKKKHASTLRQTESPPDADRTTLNHADSSKIVQHRLLSROBEL 609
Qy 663 KERARVLLLEQARBDALAKGNKNTNTATPPCNRQLSDQODEERRQLREARQLIAAR 722
Db 610 KERARVLLLEQARBDALAKGNKNTNTATPPCNRQLSDQODEERRQLREARQLIAAR 669
Qy 723 SGVKSSELPSYGEAARLAKERSKASGDENDNIEIDTNEIIPGFVVGSGDELTYLLENDL 782
Db 670 SGVKSSELPSYGEAARLAKERSKASGDENDNIEIDTNEIIPGFVVGSGDELTYLLENDL 729
Qy 783 DTPSONSKLVDLKIKLLLEVOPOVANSPSSAAQKAVTESSEBODMSGTEDLRTERLQKTT 842
Db 730 DTPSONSKLVDLKIKLLLEVOPOVANSPSSAAQKAVTESSEBODMSGTEDLRTERLQKTT 789
Qy 843 ERPRNPVPSKOSTYVKTQLOSPSOYINRPRMKQRSIOEDTKKGNBEKAAITETQKRP 902
Db 790 ERPRNPVPSKOSTYVKTQLOSPSOYINRPRMKQRSIOEDTKKGNBEKAAITETQKRP 849
Qy 903 SEDEVLANKGFQDSQYVGEALALENEBOKIDTRALVEKRLRYLMDTGRNTEBEAAMQ 962
Db 850 SEDEVLANKGFQDSQYVGEALALENEBOKIDTRALVEKRLRYLMDTGRNTEBEAAMQ 909
Qy 963 EWFMLVNNKNAIIRRMQSLLEKEHDLERRYELNRELAMLAIEDWQTEAQRERQL 1022
Db 910 EWFMLVNNKNAIIRRMQSLLEKEHDLERRYELNRELAMLAIEDWQTEAQRERQL 969
Qy 1023 LLDLVALVNRDVALVRDLDAQEKQAEDEDEHLERTLEQNGKAKKEKCVLQ 1076
Db 970 LLDLVALVNRDVALVRDLDAQEKQAEDEDEHLERTLEQNGKAKKEKCVLQ 1023

```

## RESULT 3

```

US-10-408-765A-1942
; Sequence 1942, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Boia D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Marnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; CURRENT APPLICATION NUMBER: US/10/408, 765A
; NUMBER OF SEQ ID NOS: 3073
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1942
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1942

```

```

Query Match 81.7%; Score 4485; DB 4; Length 882;
Best Local Similarity 100.0%; Pred. No. 1.5e-23;
Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 195 MSKTLVADSSKTEBEEDLDESPPYEPKSTPPPNLVNPFVQLETERRYKRAAPPVLPSP 254
Db 1 MSKTLVADSSKTEBEEDLDESPPYEPKSTPPPNLVNPFVQLETERRYKRAAPPVLPSP 60
Qy 255 KTGVLNENTVAGAGDLSLSPKSPVLPSPVLRGKKNASQSLVWCKEYVQYRGYKTNPT 314
Db 61 KTGVLNENTVAGAGDLSLSPKSPVLPSPVLRGKKNASQSLVWCKEYVQYRGYKTNPT 120
Qy 315 TSWNGLSFCAILHFRPDLIDYKSLNPQDIKENNKAAYDGFASIGISRLLEPSDVLIA 374
Db 121 TSWNGLSFCAILHFRPDLIDYKSLNPQDIKENNKAAYDGFASIGISRLLEPSDVLIA 180

```

```

Qy 375 IPDKLTWVTYLYQIRAHFSGGELNVQIEENSSKSTYKGVNVEYEDTNNSSVDQEKFYAELS 434
Db 181 IPDKLTWVTYLYQIRAHFSGGELNVQIEENSSKSTYKGVNVEYEDTNNSSVDQEKFYAELS 240
Qy 435 DLKRPBELQPIISGAVDFLSQDSSVFVNDSCVGESESEHQTPDDHLSPTASPYCRRTKS 494
Db 241 DLKRPBELQPIISGAVDFLSQDSSVFVNDSCVGESESEHQTPDDHLSPTASPYCRRTKS 300
Qy 495 DTEPQKSGQSSGRSSGSDPGICNTDSTQAVLLGKRLKAEFLKESDLYVSPKKDM 554
Db 301 DTEPQKSGQSSGRSSGSDPGICNTDSTQAVLLGKRLKAEFLKESDLYVSPKKDM 360
Qy 555 SPFFICEETDQKQITDIGNLEKEKLENSRSLSECRSDPSPIKTSLSPTSRLGYSYS 614
Db 361 SPFFICEETDQKQITDIGNLEKEKLENSRSLSECRSDPSPIKTSLSPTSRLGYSYS 420
Qy 615 RDLDLAKKGHASLQTESDPDADRRTTLNHADSSKIVQHRLLSROBELKERARVLLLEQAR 674
Db 421 RDLDLAKKGHASLQTESDPDADRRTTLNHADSSKIVQHRLLSROBELKERARVLLLEQAR 480
Qy 675 RDAALKAGNKNNTATPPCNRQLSDQODEERRQLREARQLIAEASGVKMSLPSYG 734
Db 481 RDAALKAGNKNNTATPPCNRQLSDQODEERRQLREARQLIAEASGVKMSLPSYG 540
Qy 735 EMAAEKLERKSKASGDENDNIEIDTNEIIPGFVVGSGDELTYLLENDLDTPEONSKLVDL 794
Db 541 EMAAEKLERKSKASGDENDNIEIDTNEIIPGFVVGSGDELTYLLENDLDTPEONSKLVDL 600
Qy 795 KLKLLLEVOPOVANSPSSAAQKAVTESSEBODMSGTEDLRTERLQKTERPRNPVPSKD 854
Db 601 KLKLLLEVOPOVANSPSSAAQKAVTESSEBODMSGTEDLRTERLQKTERPRNPVPSKD 660
Qy 855 STYKKTQLOSPSOYINRPRMKQRSIOEDTKKGNBEKAAITETQKRPSEDEVLANKGFQD 914
Db 661 STYKKTQLOSPSOYINRPRMKQRSIOEDTKKGNBEKAAITETQKRPSEDEVLANKGFQD 720
Qy 915 TSOYVGEALALENEBOKIDTRALVEKRLRYLMDTGRNTEBEAAMQEWMLVNNKNAI 974
Db 721 TSOYVGEALALENEBOKIDTRALVEKRLRYLMDTGRNTEBEAAMQEWMLVNNKNAI 780
Qy 975 IRRNQSLLEKEHDLERRYELNRELAMLAIEDWQTEAQRERQLLDLVALVNR 1034
Db 781 IRRNQSLLEKEHDLERRYELNRELAMLAIEDWQTEAQRERQLLDLVALVNR 840
Qy 1035 DALVRDLDAQEKQAEDEDEHLERTLEQNGKAKKEKCVLQ 1076
Db 841 DALVRDLDAQEKQAEDEDEHLERTLEQNGKAKKEKCVLQ 882

```

## RESULT 4

```

US-10-450-763-53241
; Sequence 53241, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyeseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450, 763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 53241
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

```

```
; NAME/KEY: DOMAIN
; LOCATION: (61)..(91)
; OTHER INFORMATION: Actinin-type actin-binding domain proteins domain identified
; OTHER INFORMATION: by EMATRIX, accession number BL00019D, p-value=4.200e-19, raw bcc
; OTHER INFORMATION: of 15.33
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (42)..(147)
; OTHER INFORMATION: Calponin homology (CH) domain identified by Pfam, accession
; OTHER INFORMATION: name CH, E-value=2.4e-17, Pfam score of 71.1
US-10-450-763-53241
```

Query Match 57.4%; Score 3151; DB 5; Length 669;

Best Local Similarity 89.5%; Pred. No. 1e-161; Matches 634; Conservative 6; Mismatches 12; Indels 56; Gaps 4;

```
QY 248 APPVLSFKTVLNTVTSAGKDLST--PKSPSPSVLGRKPNASGLLWCKEVTKNY 305
DB 2 AKDYLYKKYG--NQGVAAIPDVPAWAEAPSPSPVLGRKPNASOSLLWCKEVTKNY 59
QY 306 RGKVTNFTTSMWNGLSFCALIHFRPDLIDYSLNPODKENKKAVIDGPASIGSRLL 365
DB 60 RGKVTNFTTSMWNGLSFCALIHFRPDLIDYSLNPODKENKKAVIDGPASIGSRLL 119
QY 366 EPEDMVLAIIDKLTWTVTYLQIRAHFSGQELNVQIEENSSKSTYKGVYETDITNSVD 425
DB 120 EPEDMVLAIIDKLTWTVTYLQIRAHFSGQELNVQIEENSSKSTYKGVYETDITNSVD 179
QY 426 QEFYVAFSLDLKREBELQPISGAVDFLSQDDSVFVNDSGVGESESHQTPDDHLSPTA 485
DB 180 QEFYVAFSLDLKREBELQPISGAVDFLSQDDSVFVNDSGVGESESHQTPDDHLSPTA 239
QY 486 SPVCRRTKSPTEPOKSGQSSGRTSGSDPGICSTVDSIOAVLLGKRLLKATLEISDL 545
DB 240 SPVCRRTKSPTEPOKSGQSSGRTSGSDPGICSTVDSIOAVLLGKRLLKATLEISDL 299
QY 546 YVADKKKDKMSPPFICEETDEQKLTLDIGSNLEKEKLENSRLECRSDPESPIKTSLS 605
DB 300 YVADKKKDKMSPPFICEETDEQKLTLDIGSNLEKEKLENSRLECRSDPESPIKTSLS 359
QY 606 TSKLGYSYSDLDLAKKGAHLRQTESDPADRTTLNHADHSSKIYOHRLLSQBELKER 665
DB 360 TSKLGYSYSDLDLAKKGAHLRQTESDPADRTTLNHADHSSKIY----- 405
QY 666 ARVLLEARDALAKAGNKNTATATPFCRQLSDQDEERRRLRARAROLIAEARS 725
DB 406 -----QDEERRRLRARAROLIAEARS 430
QY 726 KMSLPSEYGEAAAEKLEKRSKASGDENDNIEIDTNEIPEGFVVGSGDELTLNENDLTP 785
DB 431 KMSLPSEYGEAAAEKLEKRSKASGDENDNIEIDTNEIPEGFVVGSGDELTLNENDLTP 490
QY 786 EONSKLVDLKLKLEVOQVANSPPSAQAQVATESSEQDMKSGTEDLRTERLOKTERF 845
DB 491 EONSKLVDLKLKLEVOQVANSPPSAQAQVATESSEQDMKSGTEDLRTERLOKTERF 550
QY 846 RNVVVSSTKSTVKTQVQSFOYIENPPEMKRQSIQEDTRKGAEEAAATTEVRKSED 905
DB 551 RNVVVSSTKSTVKTQVQSFOYIENPPEMKRQSIQEDTRKGAEEAAATTEVRKSED 610
QY 906 EVLNKPKTSQVYVGEALALENEQKIDTRALVERGLRYLMDTGN 953
DB 611 E---KGFKTSQVYVGEALALENEQKIDTRALVERGLRYLMDTGN 655
```

RESULT 5  
US-10-450-763-53239

; Sequence 53239, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US

```
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ. ID NOS: 60736
; SOFTWARE: Cn3com
; SEQ ID NO 53239
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (65)..(97)
; OTHER INFORMATION: REV protein (anti-repression transactivator protein) domain
; OTHER INFORMATION: identified by EMATRIX, accession number PF00424A, p-value=9.521e-
; OTHER INFORMATION: 09, raw score of 14.34
US-10-450-763-53239
```

Query Match 19.0%; Score 1045; DB 5; Length 221;

Best Local Similarity 84.4%; Pred. No. 1.e-48; Matches 217; Conservative 0; Mismatches 4; Indels 36; Gaps 1;

```
QY 628 RQTESPDDADRTTLNHADHSSKIYOHRLLSROBELKERAVLLEQARRDAALRAGNKNT 687
DB 1 RNSSPDDADRTTLNHADHSSKIYOHRLLSROBELKERAVLLEQARRDAALRAGNKNT 60
QY 688 NTATPPCNRLSDQDEERRRLRARAROLIAEARSQVKNSELPYGEAAAEKLEKRSKA 747
DB 61 NTATPPCNRLSDQDEERRRLRARAROLIAEARSQVKNSELPYGEAAAEKLEKRSKA 120
QY 748 SGENDNIEIDTNEIPEGFVVGSGDELTLNENDLTPENSKLVDLKLKLEVOQV 807
DB 121 SG-----EONSKLVDLKLKLEVOQV 144
QY 808 NSPSSAAQKAVTSSBQDMKSGTEDLRTERLOKTERFRPVVFSKDSVTRKTQLOFSQ 867
DB 145 NSPSSAAQKAVTSSBQDMKSGTEDLRTERLOKTERFRPVVFSKDSVTRKTQLOFSQ 204
QY 868 YIENRPEMKRQSIQED 884
DB 205 YIENRPEMKRQSIQED 221
```

RESULT 6

US-11-097-143-39276

; Sequence 39276, Application US/11097143  
; Publication No. US20050208558A1  
; GENERAL INFORMATION:  
; APPLICANT: Ventec, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; FILE REFERENCE: DROSOPHILA GENES.  
; FILE REFERENCE: C1000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: 60/157,832  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: 60/160,191  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 60/161,932  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: 60/164,769  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/173,383  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/175,693  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: 60/184,831





QY 231 NPVOLETERBVRKAPAPVLSPKTGVNLNENTVSAGKDLSTSPK----- 275  
DB 274 NPVOLETERBVRKAPAPVLSPKTGVNLNENTVSAGKDLSTSPKCTHVTLPBRHCWK 333  
QY 276 -----PSPISP 282  
DB 334 HMTDPAITLPLP 345

RESULT 8  
US-10-473-576-5  
Sequence 5, Application US/10473576  
Publication No. US20040101884A1  
GENERAL INFORMATION:  
APPLICANT: INCYTE CORPORATION  
APPLICANT: LU, DYUNG AINA M.  
APPLICANT: KAVITZU, CHANDRA S.  
APPLICANT: GANDHI, AMEENA R.  
APPLICANT: HAFALIA, APRIL J.A.  
APPLICANT: DING, LI  
APPLICANT: LU, YAN  
APPLICANT: RAMKUMAR, JAYALAXMI  
APPLICANT: SWARNAKAR, ANITA  
APPLICANT: TANG, Y. TOM  
APPLICANT: YUE, HENRY  
APPLICANT: TRAN, BAO  
APPLICANT: LEE, SOO YUEN  
APPLICANT: WARREN, BRIDGET A.  
APPLICANT: NGUYEN, DANNIEL B.  
APPLICANT: THANGAVELU, KAVITHA  
APPLICANT: YAO, MONIQUE G.  
APPLICANT: ELLIOTT, VICKI S.  
APPLICANT: BAUGHN, MARIAH R.  
APPLICANT: EMERLING, BROOKE M.  
APPLICANT: LAL, PREETI G.  
APPLICANT: GIEZEN, KIMBERLY J.  
APPLICANT: BECHTA, SHANYA D.  
APPLICANT: MARQUIS, JOSEPH P.  
APPLICANT: KABLE, AMY B.  
TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT  
FILE REFERENCE: PP-0921 USN  
CURRENT APPLICATION NUMBER: US/10/473,576  
PRIOR FILING DATE: 2003-09-29  
PRIOR APPLICATION NUMBER: PCT/US02/09809  
PRIOR FILING DATE: 2002-03-29  
PRIOR APPLICATION NUMBER: US 60/280,387  
PRIOR FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: US 60/282,335  
PRIOR FILING DATE: 2001-04-05  
PRIOR APPLICATION NUMBER: US 60/286,663  
PRIOR FILING DATE: 2001-04-13  
PRIOR APPLICATION NUMBER: US 60/285,484  
PRIOR FILING DATE: 2001-04-19  
PRIOR APPLICATION NUMBER: US 60/350,702  
PRIOR FILING DATE: 2002-01-18  
PRIOR APPLICATION NUMBER: US 60/351,749  
PRIOR FILING DATE: 2002-01-25  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: PERL Program  
SEQ ID NO 5  
LENGTH: 1523  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incyte ID No: 1510943CD1  
US-10-473-576-5

Query Match 17 2%; Score 945.5; DB 4; Length 1523;  
Best Local Similarity 30.9%; Pctd. No. 3e-42;  
Matches 288; Conservative 84; Mismatches 147; Indels 345; Gaps 16;  
QY 266 AGKDLSTSPK-----PSPISPVLGRKPNASGLLWVCKEVTGNVGVKLTNFTTSM 317

DB 1010 AGAFRASPEKAEEDRRRLPGSOAPAL-----VSSQGLLEWCEVTTGYGVRTTNFTTSM 1066  
QY 318 RNLGSCATLHHRPRLIDYKSLNIPDINKNNKKAADGSAISGISLLEPSDMVLLAID 377  
DB 1067 RNLGAFCAILHRRYPDKIDYASLDPLNINKNNKQAADGSAISGISLLEPSDMVLLAID 1126  
QY 378 KLTVMYLYOIRAHFSGQELNVQVLENSGSKTYKGVNTEPTDINSVD-----QEKFYAEL 433  
DB 1127 KLTVMYLYOIRAHFSGQELNVQVLENSGSKTYKGVNTEPTDINSVD-----QEKFYAEL 433  
QY 434 SDIKREP-ELQOPISGAVDFLQODSVFVNDGSGVSSSEHQTDPDHLSPSTASPYCRBT 492  
DB 1187 AEGQEPKEAADRADGAAPGAVASRNVAAGRASDGGABA----- 1225  
QY 493 KSDTEPQSQ-----QSSGRTSGSDPDGICSTNDSTQAVLLGKKRLKAELELSPLYVS 548  
DB 1226 -----PRESRPAEVPAPGAVNAGAGAG-----GGVRLRRPSV 1258  
QY 549 DKKDMSPPFICETDEQKLTLDIGSNLEKEKLENSRSLGCRSDPESP1KKTSLSP7SK 608  
DB 1259 NGEPSGVP----- 1271  
QY 609 LGYSYSRDLDLAKKGAASLRQTES-----DPDADRTTLNADHSSKIYQHRLLSROBELKE 664  
DB 1272 GSFSSHVRDADLLKKRSRLRNSSSFEMDDPA----- 1303  
QY 665 RARVLLGQARRDALKAQNKHTNTATPFCNRQLSQDDEERRQLREARQLIAEABSG 724  
DB 1304 ----- 1303  
QY 725 VKMSLEPYGEMAAEKLEKRSKASGENDNIEIDTNEBIEPGFVVGGBDELNTLBNLDLT 784  
DB 1304 -----GAMGA-----AAEGAPD----- 1317  
QY 785 PQNSKLVDLKULLEVPQVANSPPSSAQAATVTESSQDMKSGTEDLRTLRLOKTER 844  
DB 1318 -----PSPAPGPPTA----- 1328  
QY 845 FRNPVVPKSDSTVVRKTQLQSFQYIENRPMRGORSIOEDTKKGNDEKAAITETOKPSE 904  
DB 1329 -----DSOOPGSGSPSEBPPSPGB 1349  
QY 905 DEVLANKFQDTSQYVYVGELEALENEBOKOIDTPAALVEKRYLIMPTGRTEEBEAMQOM 964  
DB 1350 EAGLGR-FQDTSQYVYVGELEALENEBOKOIDTPAALVEKRYLIMPTGRTEEBEAMQOM 1408  
QY 965 FVLVKKKALIRRMQQLSLLEKEHDLERRYELLNBLRAMLAIEDWQKTEAQKREBOLL 1024  
DB 1409 FVLVKKKALIRRMQQLSLLEKEHDLERRYELLNBLRAMLAIEDWQKTEAQKREBOLL 1468  
QY 1025 DELVALVKKKALIRRMQQLSLLEKEHDLERRYELLNBLRAMLAIEDWQKTEAQKREBOLL 1075  
DB 1469 EELVSLVQDDELVDLDHKEKRIALEBEDRLERGLBOERRKLSROLSSRRRCVL 1522

RESULT 9  
US-09-864-761-42973  
Sequence 42978, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Hanzel, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecmca-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456

```

; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42978
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007629.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.79
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.76
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.1
; OTHER INFORMATION: EST HUMAN HIT: BF307519.1, EVALU8 2.00e-41
; OTHER INFORMATION: SWISSPROT HIT: P32583, EVALU8 8.10e-02
; US-09-864-761-42978

Query Match      16.3%; Score 897; DB 3; Length 175;
Best Local Similarity 100.0%; Pred. No. 8.1e-41;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 352 AYDGFASIGISRLLEPSDMVLLAIPDKLTWTYTYQIRAHFSGGELNVVQIEENSSKSTY 411
DB 1 AYDGFASIGISRLLEPSDMVLLAIPDKLTWTYTYQIRAHFSGGELNVVQIEENSSKSTY 60
QY 412 KVGNYETDNTSSVQEKFYAELSDIKREPELQOPISGAVNDPLSDDDSVFVNDGSGESSES 471
DB 61 KVGNYETDNTSSVQEKFYAELSDIKREPELQOPISGAVNDPLSDDDSVFVNDGSGESSES 120
QY 472 EHQTDDHLSPTSPYCRRTKSDTEPQKSOSSGRTSGSDDPGICNTDSTQAO 526
DB 121 EHQTDDHLSPTSPYCRRTKSDTEPQKSOSSGRTSGSDDPGICNTDSTQAO 175

RESULT 10
US-10-029-386-33831
; Sequence 33831, Application US/10029386
```

```

; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AROMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33831
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007098.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.61
; OTHER INFORMATION: SWISSPROT HIT: P32583, EVALU8 1.10e-01
; US-10-029-386-33831

Query Match      16.3%; Score 897; DB 4; Length 175;
Best Local Similarity 100.0%; Pred. No. 8.1e-41;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 352 AYDGFASIGISRLLEPSDMVLLAIPDKLTWTYTYQIRAHFSGGELNVVQIEENSSKSTY 411
DB 1 AYDGFASIGISRLLEPSDMVLLAIPDKLTWTYTYQIRAHFSGGELNVVQIEENSSKSTY 60
QY 412 KVGNYETDNTSSVQEKFYAELSDIKREPELQOPISGAVNDPLSDDDSVFVNDGSGESSES 471
DB 61 KVGNYETDNTSSVQEKFYAELSDIKREPELQOPISGAVNDPLSDDDSVFVNDGSGESSES 120
QY 472 EHQTDDHLSPTSPYCRRTKSDTEPQKSOSSGRTSGSDDPGICNTDSTQAO 526
DB 121 EHQTDDHLSPTSPYCRRTKSDTEPQKSOSSGRTSGSDDPGICNTDSTQAO 175

RESULT 11
US-11-097-143-23565
; Sequence 23565, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CU000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23565
; LENGTH: 996
; TYPE: PRT
```



Qy 698 LSDQDEERRRQRRARQOLIAEARSQVMSSELPSEYGAEMAEKLEKERSKASG--DENUNI 755  
Db 437 LPKQMD-----LESPKQVLE-----TKASETDYEDLNPDKODSSKGANPPDSSD- 484  
Qy 756 EIDNBEIRP---EGFVVGSGDE-----LTINENDLTPPEONSKLVDLKTKLEFVQOV 806  
Db 485 EVELIKALPAOOSKQKVPVPPPPPKIGISISINSEKPHSSPTL--SHGKKMPPTPRI 542  
Qy 807 ANSPSSAAQAVTESSEODMSKGTEDLTERLQ--KTERF-----RNP 848  
Db 543 SIKTQTPAKMT---HOGQKSSISSSSSEHLNSTRDRGADRGSISLPSANGPKP 599  
Qy 849 VVFSKSTVRK-----TOLQSFQYIENRPEMKORS-----IOED-----TKKXNEK 892  
Db 600 LRASVGSPLRSEBSPTLSLSITS-----PWKKRQAPLPPIQTPDPSDGFSLSDQ 654  
Qy 893 AAITETQK-----PSEDEVANKGFKOTQY----- 918  
Db 655 KALHTOLKAPNLGDSRRLPLDQSLSDSEATESNYESLSISNADEVVYRRLV 714  
Qy 919 ----- 918  
Db 715 PPTQPTENTVERSKEDQSPVIYNDFDNVSPGLHNSKTHGKRRKKGAPAVPIPPKVL 774  
Qy 919 -----VGEIALLENQKQIDTRPALVEKRLR-----YLMDTG-----RNTBE 956  
Db 775 QRLPLQEIHRHEFELIIVQQLQLEKQVLEKMIHRCRSLDADYDTPDSEAEVLVNSKE 834  
Qy 957 BEAMQEWMLVNKNALIRPMQSLSEKEDLERRYELLNREIRAMLAIEDMQTEAQ 1016  
Db 835 VEDTILQLELVNKEKLEFRQAEIMYLRQHRLEQOADIETHEIRVLMQPEHNKTSD 894  
Qy 1017 KRREQLLDELVALVNRKDALVRLDQAEKQAEEDHELE-----RTL 1059  
Db 895 KAHSEVVLNRLVKNVEMNEVIDSLETDVRBARDSIKNRLHIYNSEREPAPHPSA 954  
Qy 1060 EQNKGMAKKEK 1072  
Db 955 DKSSKLSKKERK 967

RESULT 13  
US-10-359-012-2  
; Sequence 2, Application US/10359012  
; Publication No. US20030232419A1  
; GENERAL INFORMATION:  
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
; APPLICANT: KOLODKIN, Alex L.  
; APPLICANT: TERMAN, Jon R.  
; APPLICANT: MAO, Tianyi  
; APPLICANT: PASTERKAMP, Ronald J.  
; APPLICANT: YU, Hung-Hsiang  
; TITLE OF INVENTION: MOLECULES INTERACTING WITH CASL (MICAL) POLYNUCLEOTIDES, POLYPEPT  
; FILE REFERENCE: JHU1840-3  
; CURRENT APPLICATION NUMBER: US/10/359,012  
; CURRENT FILING DATE: 2003-02-04  
; PRIOR APPLICATION NUMBER: US 60/388,325  
; PRIOR FILING DATE: 2002-06-13  
; PRIOR APPLICATION NUMBER: US 60/384,302  
; PRIOR FILING DATE: 2002-05-30  
; PRIOR APPLICATION NUMBER: US 60/354,178  
; PRIOR FILING DATE: 2002-02-04  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 863  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-359-012-2

Query Match 9.1%; Score 498.5; DB 4; Length 863;  
Best Local Similarity 24.9%; Pred. No. 1.9e-18;

Matches 218; Conservative 137; Mismatches 367; Indels 153; Gaps 30;  
Qy 293 SLVWCKEVTNKNVGVKITNTTSMRNGLSICALTHHFRPDLIDYKSLNPDIDENKKA 352  
Db 7 ALLAMCRQCGYGVIEIRDLSSFRDGLACALHRRPDLPLDLSIKONVENNRLA 66  
Qy 353 YD-GFASIGISRLPESDMVLAIPLDKLTWVTVYQIRAHPSG----- 394  
Db 67 FEVAEKELGIPALDPRDMWSMVPDCLSTITTYSQYNNHCSGQAGVSPRKLAPCS 126  
Qy 395 -----QELNVQIEBNSKSTYKGVNEDTNS-----VDQEKFYA 431  
Db 127 PPSVAPTPVEBEDVAQGEELSSGSLSEQGTQTPSSCAACQGVHLYQRYLAGRLYHR 186  
Qy 432 ELSDLKEPELQOISGAVDFLSODDSVFVNVDSGVSESHQCPDDHLSTASPYCR 491  
Db 187 HCFRCRCSSTLPL--GAYE--NGPESETPVC-----ABHCA--RLGSGTSG--TR 230  
Qy 492 TKSDEPQKSGQSGRTSGSDPGICENTDSTQAVLLGKRLKAELELSDLVYSDK 551  
Db 231 PGFSGQKQHQOQADADADVPGCGPSSAPAGAEADGPASPEARQPTKRVPKL 290  
Qy 552 KDM-SPPF-----ICEETDEQKLTLDIGSNLEKCL--ENSSSLECRSDPESP 599  
Db 291 QELASPAGPPTPAPRKAESTTTPAPPTPRPSLQENLVQAGSSSLVNGRLHELVP 350  
Qy 600 KTSLSPTSKGYSTRD-----LDA-----KKKASLQTES--DPDADRTTLNADHSSK 649  
Db 351 KPRGTPPSEGTAPAPRDPPWITLVQAEPKKPPPLPSSPPGSDSRQVENGG--TEE 409  
Qy 650 IVQ-----HRLSROELKERARVLLBOARDALAKANKNT--NTATPFCNRQ 697  
Db 410 VAQSPPTASLESKRYNPFEEHEDEKHEAPASLALSPAL--GHPESTPSLHPWGIT 467  
Qy 698 LSDQDEERRRQRRARQOLIAEARSQVMSSELPSEYGAEMAEKLEKERSKASGENDNIEI 757  
Db 468 PTSSPKTKQPPAPAPASPLALASRLSHSEPPSATPPPALSVESLSSESASQTAGEL 527  
Qy 758 DTNEIIPGFGVVGSGDELTVLENDLTPPEONSKLVDLKTKLE-----VQPVANSFSS 812  
Db 528 LEPVAVPK-----SSSEPAVHAPGTPGNFVSLSTNSSLSAGELVPEPVEQWFOA 577  
Qy 813 A---AOKAVTESSEODMSKGTEDLTERLQKTERFPNPFVFSKSTVRKQLOQSFQYI 869  
Db 578 SGLAPRTSSSGGQPAKPCSGATPTLLVGD--RSPVPSGESS--POLVYKSSCK 631  
Qy 870 ENRPEMKORSIOEDTKKGBEKAITEYORKSEDEVLNKGFK-----DTSQYVV--- 920  
Db 632 ENPFRKPSAPASPATKATKS-----KVPAPAPGHPFLIRKTVQADYIIBED 683  
Qy 921 --GELALLENQKQIDTRPALVEKRLYLMDTGRNTEBEAMQEWMLVNKNALIRRM 978  
Db 684 IHGMDTIERRLDLERGVLLBEKRLGGLNEG--EDMDLVDMFKLTHEKILVARE 739  
Qy 979 NQSLSEKEDLERRYELLNREIRAML--AIEDMQTEAKRREQLLDELVALVNRDA 1036  
Db 740 SELTVFKQQLRQADVVEIRCLNKPKED--TEBPRARKVLMQELVTLIBORNA 797  
Qy 1037 LVRLDQAEKQAESEDEHLERTLEQNGKAKTEE 1071  
Db 798 IINCLDEDRQREBEDQMLE-----AMIKKKE 824

RESULT 14  
US-10-359-012-14  
; Sequence 14, Application US/10359012  
; Publication No. US20030232419A1  
; GENERAL INFORMATION:  
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
; APPLICANT: KOLODKIN, Alex L.  
; APPLICANT: TERMAN, Jon R.  
; APPLICANT: MAO, Tianyi  
; APPLICANT: PASTERKAMP, Ronald J.

APPLICANT: YU, Hung-Hsiang  
TITLE OF INVENTION: MOLECULES INTERACTING WITH CASL (MICAL) POLYNUCLEOTIDES, POLYPEPT  
FILE REFERENCE: JHU1840-3  
CURRENT APPLICATION NUMBER: US/10/359,012  
CURRENT FILING DATE: 2003-02-04  
PRIOR APPLICATION NUMBER: US 60/388,325  
PRIOR FILING DATE: 2002-06-13  
PRIOR APPLICATION NUMBER: US 60/384,302  
PRIOR FILING DATE: 2002-05-30  
PRIOR APPLICATION NUMBER: US 60/354,178  
PRIOR FILING DATE: 2002-02-04  
NUMBER OF SEQ ID NOS: 40  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 14  
LENGTH: 863  
TYPE: PR1  
ORGANISM: Homo sapiens  
US-10-359-012-14

Query Match 9.1%; Score 498.5; DB 4; Length 863;  
Best Local Similarity 24.9%; Pred. No. 1.9e-18;  
Matches 218; Conservative 137; Mismatches 367; Indels 153; Gaps 30;

293 SLVWCKEVTKNYRGVKTINFTTSMWNGSLFCAILHHRPDLIDYSLNPQDIKENNKA 352  
7 ALLMCRQCEGYGVIRDLSSFRDGLAFCAILHHRPDLIDPDSLNDVFNENRRLA 66  
353 YD-GFASIGISRLLEPSDMLAIPLDKLTWYTYIYQIRAHFSG----- 394  
67 FEVAEKELGIPALDDNDVMSVPCDLSIMTYVSQYTNHFCSPQAGVSPRKGILAPCS 126  
395 -----OELNVQIEENSSKSTYKGVNGETDINS-----VDQEKFYA 431  
127 PPSVAPTPVEBEDVAQGEELSSGLSEGTGTPTSSCAACQOHVHLYQRYLADGRLYHR 186  
432 ELSDLRREPLQPIGAVDFLSQDDSVFVNDGVSSESHQTPDDHLSPTASPYCR 491  
187 HCFRCRCSTLLP--GAYE-NGPEEGTFVC-----AEHCA--RLGPGTRSG--TR 230  
492 TKSCTEPOKQSGSGRTSGDDPGICSNITDSTQAVLIGKKLLKATLELSDLYSDK 551  
231 PGFSPQKQOHOQOLABDAKVPGGGPSSAPAGAADGPASPEARPOIPTYKRVPGKL 290  
552 KDM-SPPF-----ICEETDEQKLTLDIGSNTLEKEKL---ENSRLECRSDPESPIK 599  
291 QELASPPAGRPAPAKASSTTPAPTPRPSLSQOENLVQAGSSSLVNGRLHLPVP 350  
600 KTSLSFTSKLGYSYSD--LDLA---KKHNASLRQTES--DPDADRTLLNHADHSK 649  
351 KPGTPEKPSGTPAPRKDPWITLVQAEPKKPAPLPPSSPGPPODSRQVENGG--TEE 409  
650 IYQ-----HLLSROEELKERARVLLQARDAALKAGNKNT--NTATPPCNQ 697  
410 VQAPSPFTASLESKPYNPFEEBEDKEEAPAPASLATSPAL--GHPESTPKSLHPWYGIT 467  
698 LSPQDEERRQLREARQLIAEARGVMSSELPSYGEAAEKLKERSKASGDENDNIBI 757  
468 PTPSPPTKGRPARASASPLALHASTLSHSEPPSATPSALVESLSSSASQITGATL 527  
758 DTNEERIEGFFVVGDELITNLENDLTPQNSKLVLDLKKLLE-----VQOPVANSPPS 812  
528 LEPFAVPK-----SSSEPAVNAFGTPGNPVSLSSTNSLASSELVPERVEQMPOA 577  
813 A---AQKAVTESSEQOMKSGTEDLRTBLQKTERFRNPVYVESKDSVTWKTOLOQSPQYI 869  
578 SPGLARTSGSSGPQAKPCSGATPTPLLVGD--RSPVSPGSSS--PQLOVYSSCK 631  
870 ENRPEKRORSIGEDTYKGNBEKAAITETQKPSDEVLNKGPK-----DTSQYVV--- 920  
632 ENPFRPKRPPAASPAIKKATKS-----KVPRPAPGHPPLIKRKVQADQYIPEED 663  
921 --GELALLENBOKQIDTRALVYKRLRLYIMDTGRNTEEBEAMQEWFMVLNKKNALIRRM 978

684 IHGMDTIERRLDALBHRGVLTSEKRLGGLNEGR-----EDMDLVDFKLIHEKLLVRR 739  
979 NQSLSEKENDLRRRYLLNRELPAH--AIDMQTKAQRREQLDLVLVNRDA 1036  
740 SELLVFKQOOLORQADVYELRCLLNKEKWM--TEEDRARKVLMQBLVTLIEGRNA 797  
1037 LVRLDQAEKQAEDEHLERTLEQNGKMAKKE 1071  
798 IINVLDEDRQREBEDKMLE-----AMIKKE 824

RESULT 15  
US-11-169-041-167  
Sequence 167, Application US/11169041  
Publication No. US20060019284A1  
GENERAL INFORMATION:  
APPLICANT: Bristol-Myers Squibb Company  
TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES FOR PREDICTING ACTIVITY OF  
TITLE OF INVENTION: COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE  
TITLE OF INVENTION: KINASES AND/OR PROTEIN TYROSINE KINASE PATHWAYS IN LUNG CANCER  
FILE REFERENCE: 10001 NP  
CURRENT APPLICATION NUMBER: US/11/169,041  
CURRENT FILING DATE: 2005-06-28  
PRIOR APPLICATION NUMBER: 60/584,405  
PRIOR FILING DATE: 2004-06-30  
NUMBER OF SEQ ID NOS: 527  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 167  
LENGTH: 863  
TYPE: PR1  
ORGANISM: Homo sapiens  
US-11-169-041-167

Query Match 9.1%; Score 498.5; DB 6; Length 863;  
Best Local Similarity 24.9%; Pred. No. 1.9e-18;  
Matches 218; Conservative 137; Mismatches 367; Indels 153; Gaps 30;

293 SLVWCKEVTKNYRGVKTINFTTSMWNGSLFCAILHHRPDLIDYSLNPQDIKENNKA 352  
7 ALLMCRQCEGYGVIRDLSSFRDGLAFCAILHHRPDLIDPDSLNDVFNENRRLA 66  
353 YD-GFASIGISRLLEPSDMLAIPLDKLTWYTYIYQIRAHFSG----- 394  
67 FEVAEKELGIPALDDNDVMSVPCDLSIMTYVSQYTNHFCSPQAGVSPRKGILAPCS 126  
395 -----OELNVQIEENSSKSTYKGVNGETDINS-----VDQEKFYA 431  
127 PPSVAPTPVEBEDVAQGEELSSGLSEGTGTPTSSCAACQOHVHLYQRYLADGRLYHR 186  
432 ELSDLRREPLQPIGAVDFLSQDDSVFVNDGVSSESHQTPDDHLSPTASPYCR 491  
187 HCFRCRCSTLLP--GAYE-NGPEEGTFVC-----AEHCA--RLGPGTRSG--TR 230  
492 TKSCTEPOKQSGSGRTSGDDPGICSNITDSTQAVLIGKKLLKATLELSDLYSDK 551  
231 PGFSPQKQOHOQOLABDAKVPGGGPSSAPAGAADGPASPEARPOIPTYKRVPGKL 290  
552 KDM-SPPF-----ICEETDEQKLTLDIGSNTLEKEKL---ENSRLECRSDPESPIK 599  
291 QELASPPAGRPAPAKASSTTPAPTPRPSLSQOENLVQAGSSSLVNGRLHLPVP 350  
600 KTSLSFTSKLGYSYSD--LDLA---KKHNASLRQTES--DPDADRTLLNHADHSK 649  
351 KPGTPEKPSGTPAPRKDPWITLVQAEPKKPAPLPPSSPGPPODSRQVENGG--TEE 409  
650 IYQ-----HLLSROEELKERARVLLQARDAALKAGNKNT--NTATPPCNQ 697  
410 VQAPSPFTASLESKPYNPFEEBEDKEEAPAPASLATSPAL--GHPESTPKSLHPWYGIT 467  
698 LSPQDEERRQLREARQLIAEARGVMSSELPSYGEAAEKLKERSKASGDENDNIBI 757  
468 PTPSPPTKGRPARASASPLALHASTLSHSEPPSATPSALVESLSSSASQITGATL 527  
758 DTNEERIEGFFVVGDELITNLENDLTPQNSKLVLDLKKLLE-----VQOPVANSPPS 812  
528 LEPFAVPK-----SSSEPAVNAFGTPGNPVSLSSTNSLASSELVPERVEQMPOA 577  
813 A---AQKAVTESSEQOMKSGTEDLRTBLQKTERFRNPVYVESKDSVTWKTOLOQSPQYI 869  
578 SPGLARTSGSSGPQAKPCSGATPTPLLVGD--RSPVSPGSSS--PQLOVYSSCK 631  
870 ENRPEKRORSIGEDTYKGNBEKAAITETQKPSDEVLNKGPK-----DTSQYVV--- 920  
632 ENPFRPKRPPAASPAIKKATKS-----KVPRPAPGHPPLIKRKVQADQYIPEED 663  
921 --GELALLENBOKQIDTRALVYKRLRLYIMDTGRNTEEBEAMQEWFMVLNKKNALIRRM 978

```

Db 468 PTSSPTKKRPPAPRAPSPLALHASRLSHSEPPSATPPALSVESLSSESASQTAGAEI 527
QY 758 DTNEIRPGFVVGGEDELJTNLENDLDTPEONSKYDLKXKLLB-----VOPQVANSPPS 812
Db 528 LEPNAVFK-----SSSEPVAHAPGTGPNPVSLSSTNSLASGELVEPRVQMPQA 577
QY 813 A--AQKAVTESSEODMKSGTEDLNTERLOKTERFRNPVVFSSKDTVKTQLOSPSQYI 869
Db 578 SPGLAPRTRGSSGCPAKPCSGATPTPLLVGD---RSPVSPGSSS---PQLQVKSCK 631
QY 870 ENRPEMKRORSIQEDTKKGNBEKAAITETOKKPSDEVINKGFK-----DTSQYV--- 920
Db 632 ENPFVRKPSPAASPATKATKGS-----KPVRRPAPGHGFPPLIKRKVQADQYIPEED 683
QY 921 --GELALENEQKOIDTEAALVEKRLRYLMDTGRTEEEBEMMOEWFMVLVNGKALLIRRM 978
Db 684 IHGEMDTTERRIDALEHKGVLLEKRLRGGLNEGR---EDDMLVDMFKLIHKHLLVRE 739
QY 979 NQLSLBEKHDLERRYEELNREILRAML--AIEDMQTEAQKRRBOLLDELVALVNRKDA 1036
Db 740 SELIYVFKQONLEQQAADVBYELRCLINKPEKDW--TEEDRAREKVMQELVTLIEQRNA 797
QY 1037 LVRLDIAQEQAAEBEDELERTLEQNKGKMAKKEE 1071
Db 798 IINCIDBDQRQBBEBEDKMLB-----AMTKKKE 824

```

Search completed: May 20, 2006, 17:30:18  
 Job time : 193 secs



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2006, 17:27:13 ; Search time 11 Seconds  
(without alignments)  
208.951 Million cell updates/sec

Title: US-09-856-723a-8  
Perfect score: 5489  
Sequence: 1 MKQYASPMPTQTVKLFKP.....RTLEQNKGMKKEKCVLQ 1076

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 21570 seqs, 2136119 residues

Total number of hits satisfying chosen parameters: 21570

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA New:  
1: /EMC\_Celerra\_SIDS3/prodata/2/pubppa/US09\_NEW\_PUB.pep:\*  
2: /EMC\_Celerra\_SIDS3/prodata/2/pubppa/US06\_NEW\_PUB.pep:\*  
3: /EMC\_Celerra\_SIDS3/prodata/2/pubppa/US07\_NEW\_PUB.pep:\*  
4: /EMC\_Celerra\_SIDS3/prodata/2/pubppa/US08\_NEW\_PUB.pep:\*  
5: /EMC\_Celerra\_SIDS3/prodata/2/pubppa/PCF\_NEW\_PUB.pep:\*  
6: /EMC\_Celerra\_SIDS3/prodata/2/pubppa/US10\_NEW\_PUB.pep:\*  
7: /EMC\_Celerra\_SIDS3/prodata/2/pubppa/US11\_NEW\_PUB.pep:\*  
8: /EMC\_Celerra\_SIDS3/prodata/2/pubppa/US60\_NEW\_PUB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	255.5	4.7	849	US-10-972-140-1	Sequence 1, Appli
2	233	4.2	2871	US-10-505-928-100	Sequence 100, App
3	223.5	4.1	3113	US-10-505-928-325	Sequence 325, App
4	222.5	4.1	915	US-10-523-014-4	Sequence 4, Appli
5	213.5	3.9	919	US-11-251-465-31	Sequence 31, Appli
6	202	3.7	2829	US-10-525-621-1	Sequence 1, Appli
7	185	3.4	3256	US-10-505-928-357	Sequence 357, App
8	184	3.4	1120	US-10-505-928-213	Sequence 213, App
9	182	3.3	1384	US-10-505-928-799	Sequence 799, App
10	180	3.3	1674	US-10-511-937-2587	Sequence 2587, App
11	167	3.0	4373	US-11-118-524-2	Sequence 2, Appli
12	166	3.0	851	US-10-511-937-2597	Sequence 2610, App
13	164.5	3.0	671	US-10-196-749-346	Sequence 346, App
14	164	3.0	1610	US-11-253-276-22	Sequence 22, Appli
15	162.5	3.0	1186	US-10-511-937-2566	Sequence 2566, App
16	162	3.0	1125	US-11-252-276-152	Sequence 152, App
17	159.5	2.9	1234	US-11-327-900-6	Sequence 6, Appli
18	157	2.9	638	US-10-511-937-2610	Sequence 2610, App
19	154.5	2.8	1531	US-10-505-928-217	Sequence 217, Appli
20	151	2.8	769	US-10-523-043-9	Sequence 9, Appli
21	150.5	2.7	730	US-10-505-928-841	Sequence 841, App
22	150	2.7	798	US-10-511-937-2451	Sequence 2451, App
23	147.5	2.7	663	US-11-169-140-14	Sequence 14, Appli
24	147.5	2.7	1809	US-10-559-415-190	Sequence 190, App
25	147.5	2.7	1919	US-10-559-415-2	Sequence 2, Appli

26	146.5	2.7	1912	US-10-511-937-2561	Sequence 2561, App
27	145.5	2.7	1056	US-10-505-928-225	Sequence 225, App
28	145.5	2.7	1212	US-10-505-928-201	Sequence 201, App
29	145	2.6	3396	US-10-505-928-449	Sequence 449, App
30	144	2.6	1575	US-10-505-928-257	Sequence 257, App
31	142.5	2.6	284	US-10-505-928-447	Sequence 447, App
32	141.5	2.6	586	US-10-505-928-114	Sequence 114, App
33	140.5	2.6	1066	US-10-511-455-2	Sequence 2, Appli
34	139.5	2.5	1066	US-10-370-959-2	Sequence 2, Appli
35	139.5	2.5	1066	US-11-302-678-50	Sequence 50, Appli
36	139	2.5	1050	US-10-505-928-347	Sequence 347, App
37	137.5	2.5	783	US-10-511-937-2985	Sequence 2985, App
38	135.5	2.5	823	US-10-511-937-2946	Sequence 2946, App
39	132	2.4	1395	US-10-505-928-513	Sequence 513, App
40	129.5	2.4	2351	US-11-183-218-310	Sequence 310, Appli
41	128	2.3	284	US-10-505-928-670	Sequence 670, App
42	128	2.3	1193	US-10-505-928-537	Sequence 537, App
43	127.5	2.3	1066	US-10-505-928-801	Sequence 801, App
44	126.5	2.3	502	US-10-511-937-3008	Sequence 3008, App
45	126.5	2.3	2351	US-11-280-757-35	Sequence 35, Appli

#### ALIGNMENTS

RESULT 1  
US-10-972-140-1  
Sequence 1, Application US/10972140  
Publication No. US2006008892A1  
GENERAL INFORMATION:  
APPLICANT: Yeart, Ilona F.  
APPLICANT: Kimberly-Clark Worldwide, Inc.  
TITLE OR INVENTION: Detection of Trichomonas  
FILE REFERENCE: 1443.168051  
CURRENT APPLICATION NUMBER: US/10/972.140  
CURRENT FILING DATE: 2004-10-22  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 849  
TYPE: PRT  
ORGANISM: Trichomonas vaginalis  
US-10-972-140-1

Query Match 4.7%; Score 255.5; DB 6; Length 849;  
Best Local Similarity 20.6%; Pred. No. 5.2e-10;  
Matches 170; Conservative 152; Mismatches 286; Indels 219; Gaps 43;

QY	290	ASGSLVWCKEYKNTKGVKVTPTTSWRKGLSCALIHFRPDLIYKSLNPODIKENN	349
DB	134	ARDALLMAKKNTOGHEVAVVNNFTTSWNTGLAFAALINFRPVLIDYSLDYNDKQAC	193
QY	350	KKAVDGPASIGISLLEPSDMVLLAIPDKLTVMVTYVQIRAHFSGELNV-----	399
DB	194	EKAPACKKELGIYVLDPEB-VIDTTPDEKSVTVQVAF-PHPASASKTAAADKIKRT	251
QY	400	--VOIEENSSKSTYKGVNVTETDNTSSVDQKFFYAEISDLRREPELQOPISGAVFLSODD	457
DB	252	VAIOKQIDELKNTFYIEDAKAIEKMTVEDEKLKADYVE-KTIPGIRKGLASVISY-----	305
QY	458	SVFPNDSGVSESEHPTP--DHLSPSTAS--PYCRKTSQDTP-----OKSQSSGRTS	509
DB	306	-----NRDIRPEIVDHRRAQMSMAALVTKCSGNNRPIETFGLEPALTN	352
QY	510	GSDPGICSNWDSYQAVLLGKRLKAKETLEISLTVSDKKQMSPPFI-----CRETD	564
DB	353	KENE---IEGTSTTRDEL-----TQELNDM-IKKQVDEPMKAKMDIINKCAIH	398
QY	565	EQKQITLDISNLEK-----EKLENSKSLGCRSDPSSPIKTSLSPTSFKLGYSYRDLDL	619
DB	399	EE-VKITE-GTTAKKQKVQKCLHKAEDLOPALAELTFLPFO-----EL	439

```

QY 620 AKKKHSAQRQESDDADRTTLINADHSHKIYQHLLNSQEBELKEARVILEQARDAL 679
Db 440 VELRINTLSSQTDISVNR-----HSHQITY-----IKLLEQUNGLKEETNEARI 486
QY 680 KAGNKHNTNTATPCNROLSDQODEERRRQLEPARQOLIAEARSQVNSKELPSYGEMAA 738
Db 487 ---NEYNA-LAQPLVDEAIAFKREBVLATISGELRERTQOLA-----KQAEAPT-----KR 532
QY 739 EKAKERSQA-SGDENDNIEIDTNEBELPEGFVVGSGDELTNLENDLDTPEQNSKLYDLKAK 797
Db 533 EHVNEIDIPFQLEKDSLHLRVNH-----SPLEINNYAVTLQH---ITYEIN 577
QY 798 KLEVEQOVANSPSSAAQ-----KAVTESSEQDMKSGTEDLRTERLOKTERFRAPVVS 852
Db 578 KIFPE-EVNAFPAARAVPIIDISITLALVYSSH-----IP 609
QY 853 KDSIVRKTOLOSFSQYIENRPEMKRORSIQEDTKKNEEKAALITETORPSDEVYANKGF 912
Db 610 GDAAVAKQAVE-----ENLASLDCVRR-KDPSPGSISQARSIAQUL-----I 651
QY 913 KOTSGY--VGEELALENBQKOIDTRBALVEXRLAYIMDTGNTTEEBEAMQEMFML--- 967
Db 652 KYTTYSATGELV-----QARLDIKQITILAK-----TLEEBERKRIINNYVKADE 700
QY 968 -VNKNALIRRNQOLSLKEKHDLERRAYELNRELRAM--LAIEDQOKTEAQRREBULL 1024
Db 701 HNNENHALDGKINSVD-GELBPKRQKXVE-VREEVNAKKEKAVBELTPIYEDLEKQDLHL 758
QY 1025 -----DELVALVNKRDLALVRLDLAQ---EKOAEDEDEHL 1055
Db 759 EITSPASINITFEENLIAHI---DTLVEIDIRIAAAGLEISEBEL 802

```

	RESULT 2	
	US-10-505-928-100	
	/ Sequence 100, Application US/10505928	
	/ Publication No. US20060088532A1	
	GENERAL INFORMATION:	
	/ APPLICANT: Ludwig Institute for Cancer Research et al.	
	/ TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES	
	/ FILE REFERENCE: 28967/39178	
	/ CURRENT APPLICATION NUMBER: US/10/505,928	
	/ CURRENT FILING DATE: 2004-08-27	
	/ PRIOR APPLICATION NUMBER: US 60/363,019	
	/ PRIOR FILING DATE: 2002-03-07	
	/ NUMBER OF SEQ ID NOS: 866	
	/ SOFTWARE: PatentIn 3.2	
	/ SEQ ID NO 100	
	/ LENGTH: 2871	
	/ TYPE: PRT	
	/ ORGANISM: Homo sapiens	
	US-10-505-928-100	
QY	Query Match	4.2%, Score 233, DB 6; Length 2871;
	Best Local Similarity	18.4%; Pred. No. 1e-07;
	Matches 164; Conservative 164; Mismatches 297; Indels 266; Gaps 36,	
Db	343 QDIENKKKATDGPASIGISRLRESDDVLAIPDKLTVMYTYLYGIRAHFSGQELANVOI	402
	EDRPDOQKNVDY---QLQKAQCEKENLGMOKLSSEKAIKEKEYE-----ERLRVLILQ	1179
Db	1129 EENSRSKYTV-----KVGNVETDTJNSVPOEKFYALSDLK---REPELOQP-----ISGA	449
	EETRGRKEYEVELAKVRHRHYHEMSNL-RNKETEITINITTKTIKISIMQKEDDSKNLRNQ	1233
QY	450 VDFLS-----QDDSVFVNDGVGESESEHQTPDDHLSPSTASPYCRRTKSPTBEPQSQ	503
	LDRLSRENBDLKDIEIVRLNDSII LGATEQR-----RAAEENALQKXA--	1279
Db	504 SSGTSSGDDEPGICSNIDSTOAQVULGKRKLIAKTLELSDLVYSDDKKDMSPPTCAET	563
	-----CGSSE-----IMQKHCHLEIKLVMMQORSEDNAHF-----KOS	1312
Db	1280 -----CGSSE-----IMQKHCHLEIKLVMMQORSEDNAHF-----KOS	1312

Qy	564	DEOGLQTL-DLGSNLKKNLKNSSLEBRSPESPISIKTSLSPYKGLVYSRDLDAKK	622
Db	1313	LEBAKATIQDNKKELERLKAFQEBAKRWEYEMEL-----SKVNNNDDEILSLK-	1363
Qy	623	KHASICTESPPDADRITLNLADHSSKLVQRLLSROBE-----LKERARVILE	671
Db	1364	-----NPEETINITKTTI-----HQLTMQKEEDPSGYRAQIDNLTRKRSLSIE	1407
Qy	672	QARDDALKAGNKHNTNTATPFCNRQLSD--QQDEBERRQLREPARQLIAEARSGVKN--	727
Db	1408	EIKR---LK-----NTLQTETENLRVVEDIQOQKATGSFVSKQQLVEELRVQTCWRT	1453
Qy	728	SELPSYGMALAEKKEKSKASGDNDNTE-----ID--TNE-----	761
Db	1460	EESVRY-----KQSLDDAKTTIQDNKKELERLKQILDKETNPKCLLEBENARLQRVQYDLQ	1515
Qy	762	-----EIPGFFVVGGDDELTNLENDDLPTEQNSKLVLDKLK-----KLLEAV	802
Db	1516	KANSSATETINIKLAVQE-----QELTRLRIDYERVSQERTVDDQITTRFQNSLKEQL	1568
Qy	803	QPOVANSPPSSAAQKAVTESS-----EODMKSGTEDLRTERLQKLT--TERFRNPVFSK-	853
Db	1569	QKQVYEESILNKTKTAASEDSCKRKLEEBELGMRSSLEQAIKITTNLTLQLEQASIVYKR	1628
Qy	854	-----DSTVKTCQ-----LQSFQYIE--NRPEMKQRQSIQEDTKKNEEKA	894
Db	1629	SEDLRQQRDVLDELHKKRQKQRTQBELRLSSVEALRRQLLOEQEBSVQAHLRNHFQKA	1688
Qy	895	ITE*QRKPEDEVLANKFQKTSQYVGEALLENOKOI-----DTRPA	938
Db	1689	IEDSRSLGNSKEKIEIRBQSLTENLJTEHMLLEBELNRLRLREYDDLKRGREASDKNAT	1748
Qy	939	LVEKRLRYLMDTGNTT-----EEBAMQOE-----WFLM	967
Db	1749	ILLEIRSQLQISNNRTLEQLGLINDLOREBERNLRQIEKFCQKALEASNRLOESKNQCTQV	1808
Qy	968	VNKKNALIRRNNOJS-----LLEKENDLERYELLNRLKMLALIEDMQKTEAQKREQL	1022
Db	1809	VOEHESLVVKTKVLEQDKARQLRELEENRAKSTLTLEATRYK-----QRLECEKQQLQN	1862
Qy	1023	LDELVALVNNKRDALVRLDQAQKQAEEDDEHLERTLEOUNKNGKAKKEKC	1073
Db	1863	DLNQMKQYQSKEAIRKIBSEBRKSKSREKNSLASEIERLQALIKRIEERC	1913
RESULT 3			
US-10-505-928-325			
; Sequence 325, Application US/10505928			
; Publication No. US20060088532A1			
GENERAL INFORMATION:			
APPLICANT: Ludwig Institute for Cancer Research et al.			
TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES			
FILE REFERENCE: 28967/39178			
CURRENT APPLICATION NUMBER: US/10/505,928			
CURRENT FILING DATE: 2004-08-27			
PRIOR APPLICATION NUMBER: US 60/363,019			
PRIOR FILING DATE: 2002-03-07			
NUMBER OF SEQ ID NOS: 866			
SOFTWARE: PatentIn 3.2			
SEQ ID NO 325			
LENGTH: 3113			
TYPE: PRF			
ORGANISM: Homo sapiens			
US-10-505-928-325			
Query Match			
Best Local Similarity 19.3%; Pred. No. 4,96-07;			
Matches 237; Conservative 215; Mismatches 466; Indels 293; Gaps 53;			
Qy	41	EGKATDEDMQSLASLVSMKQADIGNT--DDEEDNED-----DDENRVQNEEYKAK	89
Db	1572	QGIKNKKEIQELBQLLSSREQBLDCLRKQYLSENEMQOQKLTSTVLTLEMESKTLAAEKQTE	1631

```

QY 90 ITELINKNPLDEAKDLATVNSNPPDDPA-----BLNP 125
Db 1632 QLSLEFVARLQLOGLSSRLGI DTEDAIQGNNSCDISKEHTSETTERPHGDVHQ 1691
QY 126 FGDDPS-----EPITETASPKRTBDSFYVNS-----YNPKKQVOT----- 161
Db 1692 ICDKAOQODLMDIEKITEYTGALKPTGSCGSEGDVTNYPGPEDEKTOGSSSECSLSLSPS 1751
QY 162 -POLYNPDEBEAFVTKDS-----POSTGRKIRIPVDSKYLYADSSKTEBEELDES 214
Db 1752 GPNALVND-----FLGNQEDHNLQRYKETSNEMLRL-----HYIEDRDRVY----- 1797
QY 215 NPFPYKSTPPPPNNLVNPOVLEET-----ERRVKKKAPAPVLSPKTGVLENTVYASAKD 269
Db 1798 -----SLNEMKELDSKLMH.QEVQJMTKIECELEKIVELIKKE-----NSD 1840
QY 270 LSTSPKSPISPVYGRKPNASQ-----SLVWCKEVTYKNGVKTNTTTSR----- 318
Db 1841 LSEKLEYPSCDHQELQREVTSEGLNSDLEMHADSSREDIGDNVAKVNDSEKFERLDVE 1900
QY 319 NGLS-----FCALILH-----HFRPDL--IDYKSLNPODIKENNKGAVDGFASIGISRLBPS 368
Db 1901 NELSRIBEKASIEHEALYLEADLEVOTBEKLEKNENKOK----- 1943
QY 369 DMVLAIPDKLTWMT-YLYQIRAHFSQGLNVVQIENSSKSYKGVNTEYEDVNSVDOE 427
Db 1944 --VIVCEBELSVLTSEBNQRLGELDTWMSKKTMTALDQLSBKWKKE--TQELBSHQSBCLH 1999
QY 428 KFYAELSLKRPBPELQPIGSAVDFLSODDS-----VFVNDSD--GVGSESEHQTPD 477
Db 2000 CIOVAEAEVKETBELQTLSSVSESLDXDKTHLOEKLOSLKESDQALSJLTCELELNQIAQ 2059
QY 478 -----DHSFSTASPYGCRKTSPTREPOKSQOSSGSGTSGSDPG-----ICSNMDSQOAVL 528
Db 2060 LNKKEKLVKESSEIQAARLSBDEY--KLNVSKALMALVKEGEPALRLSSTQEBVHOLR 2117
QY 529 LGRKRL-LKAETLESLDYSDKKKMSPPFICEBTDEQKQTLTDIGNLEKEKLENSRS 587
Db 2118 RGIEKRLVRIDEADKKQHLIAKLEKERE-----RNDLSLK-----DKVNLRLRELOMSSEN 2168
QY 588 LEC-----NSDPSPIKTSLSPTSKLGYSYRDLAKKXNASLRQTESDPDARTTL 641
Db 2169 QELVILDAENSKAEVETLKTQIIEEMARSLKVFELDLVTLRSEKEMLTQIOBKQOQSLSEL 2228
QY 642 NHAHSSKIYVHRLISROEEL--KERARVLLBQARRAALAKAG--GEMNTVATPPCNR 696
Db 2229 D-----KLSFSFKSLBEEKQAEIQIKESKTAVENTLOQLKELNEAVALCGD 2277
QY 697 QLSDOODE-----ERRRQLREBARQLIABASGVK-----MSGLPSYGEAALAKLER 744
Db 2278 QEIMKATBOSLDPPIEEHQJLRNSIBKRLARLLEADDEKQQLCVLQOLKE--SEHNDLILGR 2336
QY 745 SK-----ASGDENDNIETDNEBIEPGF-----VVGDEDELTNL 778
Db 2337 VENLERLELIARTQOEHALAEANSKGEVETILKAKIEGATQSLRGLBLDVTTIRSEKENL 2396
QY 779 ENDLDTBPQNKVDLKLKLLVQPOVANSPPSAAQAVTESSEBQDMKST--EDLRTB 836
Db 2397 TNELOQ--KEGRISLELT-----INSSPENILQKEQEKQVOMKEKSSSTAMEMLOTO 2445
QY 837 RLQKTERFRNPVVSFSDTVRKTQLOSFQYIENRPEMKQRSIQE--DTKKGK----- 889
Db 2446 -LKELINREV--ALAHNDQACAKQENLSSOVECL--ELEKQALQGLIDEAKNNTYVLOS 2500
QY 890 -----EKAATETQKPSBEDVANKGPDTSQYVVGELALAEINP-----QKQIDTRA 937
Db 2501 SVNGLIOEVEDGKQLEKDEBEIRLKQKQIODQEQ-LVSKISQYEGEGHQLWKEQNLEIRN 2559
QY 938 ALV--EKRLRYLMDTGRNTEBEEBAMQEMFVLVVKXNAL--IRRNQQLSLEKHDLBERRY 994
Db 2560 LTVLEBQCIQVLOSKNASLQDTLEVLQSSYK--NLENELELTQKADKMSFVKKVWKRAKE 2617
QY 995 ELANREILRAML--AIEDWQKTEAQKRR-----EQLLT-----DELVALVVKRDALVRDL 1041

```

```

Db 2618 TELQREHEHMAQKTAELOEELSEGENRLAGELQLLLEIRKSSKQQLKELLTENSELTKSL 2677
QY 1042 DAOEKQAEDEBHEILERTLEKNGKGMAKKEK 1072
Db 2678 DQMKDOVEKEGKVAREBETAEYQRLHBAEKK 2708

RESULT 4
US-10-523-014-4
; Sequence 4, Application US/10523014
; Publication No. US20060094101A1
; GENERAL INFORMATION:
; APPLICANT: Linn, Lih-Liang
; APPLICANT: Yannoni, Yvonne
; TITLE OF INVENTION: M22 Interacting Proteins
; FILE REFERENCE: 08702-0097-00304
; CURRENT APPLICATION NUMBER: US/10/523, 014
; PRIOR FILING DATE: 2005-02-01
; PRIOR APPLICATION NUMBER: USSN 60/400, 044
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 4
; LENGTH: 915
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-523-014-4

Query Match 4.1%; Score 222.5; DB 6; Length 915;
Best Local Similarity 24.5%; Pred. No. 9.5e-08;
Matches 100; Conservative 66; Mismatches 175; Indels 67; Gaps 15;

QY 6 SPMTQTDVLYKFKPPLSKKVSAALQPSLCIFLRBGKADDEMQSLASLVSKQADIGN 65
Db 533 APSSRGGCSIXMPEBPAEPLAAVAEANGABQTRVANKA--PEGRSPLSABELMTIEDEGV 591
QY 66 LDDPEBNEDEDERVNOBEKAAKTEL-INKLNFLD-EAEKDLATVNSNPPDDPAEEL 123
Db 592 LDKKLDQSTDEBEKLLR--AAURELRQKRDQKREKERRLOEARGRGE----- 640
QY 124 NPFGDPSBEFETETASPR-----KTEDSFYNSYNPKKAVOTPOYLNPDPPE 172
Db 641 ---GRGNATETTRRHQORAAADSASVTVKTEBLVHNSQTRARTTYE-----S 689
QY 173 AFTVTKDSPOSTYKKNIRPVDMSKYLYADSKTEBEELDESNPFPKSTPPNNILVNP 232
Db 690 SFVARSSENGSGSTW-----MQTKTFSSSSSSSKMGSI FPR-----EDQASPRAGSLA-- 736
QY 233 VQELTERRYKAKAPAPVLSPKTGVLN-----ENTVSAGKDLSTSPKSPISPV 283
Db 737 --ALEKQOAEKKEKLMQAQSLPTISASQARKAMIEKLEKGAAGS--PGGRAAVVRSTG 792
QY 284 LGRKPNAS-----QSLVWCKEVTYKNGVKTNTFTTSWRNGLSFCALIHFRPDLIDYKSL 340
Db 793 FG-VFNANSIKQMLDWCRAKTRGYEHVDIONSSSSSDGMARFALVHNPPPEAFDYGQL 851
QY 341 NPQDIKENNKKAYDGFAS-IGISRLLEPSDMVLIAPDKLTWMTYLYQ 387
Db 852 SPQRRQNPFEVAFSSAETHADCPQLDTEDMVRLREBDMCVVYTYOE 899

RESULT 5
US-11-251-465-31
; Sequence 31, Application US/11251465
; Publication No. US20060094061A1
; GENERAL INFORMATION:
; APPLICANT: Brya, Reginald
; APPLICANT: Vandeghinste, Nick
; APPLICANT: Tomme, Peter
; TITLE OF INVENTION: Molecular Targets And Methods To Identify The
; TITLE OF INVENTION: Same, Useful In The Treatment Of Joint Degenerative And

```

```

: TITLE OF INVENTION: Inflammatory Diseases
: FILE REFERENCE: P30,172-A USA
: CURRENT APPLICATION NUMBER: US/11/251,465
: CURRENT FILING DATE: 2005-10-14
: PRIOR APPLICATION NUMBER: 60/619,384
: PRIOR FILING DATE: 2004-10-15
: NUMBER OF SEQ ID NOS: 880
: SOFTWARE: PatentIn version 3.3
: SEQ ID NO 31
: LENGTH: 919
: TYPE: prt
: ORGANISM: Homo sapiens
: US-11-251-465-31

```

Query Match	3.9%	Score 213.5	DB 7	Length 919
Best Local Similarity	21.7%	Pred. No. 3.9e-07		
Matches 160	Conservative 107	Mismatches 246	Indels 225	Gaps 344

Qy	436	LKREPELOQPIGANDPFLSODDSVAVNNGSVESSEHQTPDHDLSPEIASYCCRTSD	495
Db	152	LTKRPE-----DNHTQCOLVAVVEIGT-----SEQOMAQHTOLMSTPEPLRTLSF	198
Qy	496	TEPOKSQOQSGRTSGSDDPGICSNMSTDSQAQVLLGKKRLLAETL-----	540
Db	199	TPASATAPTSGCPIPTSD-----ESIPPK---SKAILISITVSLMATPPDPKGQV	248
Qy	541	-----ELSLVYSDDKKDMSPPTICGETTBQKQLT-----DTSNLEKELENS	586
Db	249	GTCGSASATLRIAOVSPGPRD--SPAPEDSPWREVRVLAFLIPDNFSPPTSGRTSDOSVRH--	305
Qy	587	SLECRSDPESPPIKKTSLSPSTKGLGYSVRDLDIAKKGJASR-----CTESPDDR	638
Db	306	-----SPLAASSPSP--QVLAOKTS---LVAQESVVRARSRLAKKTAEIPAA--	349
Qy	639	TTLNHADHSKIYOH---RLT-----SROELERARVLLQARRDAALAKGNK--	684
Db	350	-----SGRIICHSYLERLLANVEPVKQVSEQKEPEEAEF--VAALPEVPENNGNS	400
Qy	685	--HN-----TWTAIPFCNRQ--LSDDQDEERRRLQERARQLIABARSGV	725
Db	401	WPHNDTEIANSTPNPKPAASSPETPSAQOQEAKTQDADGPREPPSARAKRSYQA----	456
Qy	726	KMSLEPSYGEWAARAKLK--ERSKASGDENDNIENIDNEIIPGFVVGSGDELITLENDLT	784
Db	457	--VSELDSEQHLDESDLOPFRSKT-----PSSPCPASKV-----RPLRTFLHT	498
Qy	785	PEONSKYLDLKLKLLLEVQPOVANSPPSSAAQAV-----TESSBDM-----	826
Db	499	VQRNQMLM-----TPISAPRSVYKSPFIKNTPLRMDPKEXERQRLNLR	542
Qy	827	--KSGTEDLRTERLOKTTERRFPNPVYFVSQDSIVKRTOLQSFQYIENRPEMRQ----	880
Db	543	RKEAEOLRQKVEDRKRRLREYVLKKEERLRIKV--LQARERVQOMKEKKKQJLEQKPAQ	601
Qy	881	IOEPTKKGNEKAAITETQKRPSEDEVINKPFOTSQYVVEELALBNEQOQIOTPRALV	940
Db	602	IDETTEAKERERLLEBKAKKK-----AAAKMEVEARRQK--	638
Qy	941	E--RLRYLMDGTGNTTEEBEAMQEWFLVYK--NALLIRMNQLSLLEKHLERR--Y	994
Db	639	EDARLRMLQO-----BEERRRHE--LLQKKKEEEOERLAKAAEAKRLAEQROERREQ	691
Qy	995	ELLNRELRLAMLAIEDMQTEAKREBOLLDELVALVYKRDALVRDIDAOEKAEEBDEH	1054
Db	692	ERREQERREOERREOERREOERQALQAEQERRRREGERLQARELEQERKALRLQEOQ	751
Qy	1055	LERTLQNGKMAKKEK 1072	
Db	752	LQRELEBK-----KKEBO 765	

```

; Sequence 1, Application US/10525621
; Publication No. US2006010041841
; GENERAL INFORMATION:
; APPLICANT: Kiyosue, Yuko
; APPLICANT: Sasaki, Hiroyuki
; APPLICANT: Tsukita, Shoichiro
; APPLICANT: Eisai Co., Ltd.
; TITLE OF INVENTION: CULTURED XENOPUS LAEVIS CELL LINES
; TITLE OF INVENTION: EXPRESSING MUTANT ADENOMATOUS POLYPOSIS COLI GENE
; FILE REFERENCE: 082368-002400US
; CURRENT APPLICATION NUMBER: US/10/525,621
; CURRENT FILING DATE: 2005-02-22
; PRIOR APPLICATION NUMBER: PCT/JP03/10434
; PRIOR FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: JP 2002-241487
; PRIOR FILING DATE: 2002-08-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2829
; TYPE: prt
; ORGANISM: Xenopus laevis
US-10-525-621-1

```

Query Match 3.7%; Score 202; DB 6; Length 2829;  
 Best Local Similarity 18.1%; Pred. No. 1.2e-05;  
 Matches 225; Conservative 150; Mismatches 433; Indels 438; Gaps 52;

```

OY      66 LDDPEBNEDEDERVNOEELAAKITELINKLNFLEDA-----EKOLAIVNSNPDPDP 119
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1134 VDDVDDDKTINFSRYSRVEEQDEBETRQNKNIKAYASEHHNGEORIDVSRKKTIDVPS 1193
OY      120 AAEINPF-----GPDSEBPTTETASPRKTEDSFYNNSYNPFKEVQTPQYLNPFDE 171
Db      1194 SAQCPSEPYSNNSKQPKPKKEQVSNNSNTPTPSNRSRQQLHPNSAQSRPLNRPKOIP 1253
OY      172 EAPTIIDSPPOQTKNIR-----PYMSKTYLADSSKTEBEELD-----ESN 215
Db      1254 -----NKPPSTINQETIOTYCVEDPTICPSRQSSLISSASBDETEGRENRSGQESN 1305
OY      216 ---PFVEPK--STEPNNLVNVOELETEERRVVR-----X 245
Db      1306 NTLQITPKEISAVSKXGAVNERTRSSVHHRTNRNRLQTSNISPSDSSRAKSVBFSGAK 1365
OY      246 AP-----APP-----VLSPTGYLNNYTS-- 265
Db      1366 SPKSPKSAQTPKSPPEHYVQETPLPMFSCTSGSSLSDFESHSIASSIASSVASEHMIISGII 1425
OY      266 AGKULSTSP-----KESPIPSPLGRKPNAOSGLVWCKEVTKYNGVITLPTTS 316
Db      1426 SPSSLRPSPOQTMPSPRSKTPPRPQYQAKQDS-----KPIYDEBRGKVAK--TA 1475
OY      317 WRNGLSFCALIH-----HFRPDLIDYKSLNPQDIKENNKKAYDGF--ASIGISRLLEP 367
Db      1476 VHSATIGRVQVLOEADTLNHPA-----TSTPDGSCASSLSALSIDEP 1518
OY      368 ---SDWVLAIIPDLITMYTLYLOIRAHFSGQELNVQIE-----ENSFKSYKKYNGYETDT 420
Db      1519 YIQKDVOLKATMPVYL-----ENDQGNKAPPEKEFIDNKAKEDKRSBOEKDM 1565
OY      421 NSSVDOEKFYAE--ISDLKREP-----ELQOPISGA----- 449
Db      1566 LDPTDDIDILBECIISAMPKRKSRKNNKKVYQPTPGKPPRPVARKPSQLPVYGLLSSQNR 1625
OY      450 -----VDPLSQDD-----SVVANDSGVGESESEHQTPDDHLSPSTAGP 487
Db      1626 LQTOKHNFNHSDMPRPVYCVBETPINFSYATSLSDLTIESPSE-----PNDQOP 1676
OY      488 YCRRTKEDTEPOKSQOSGRTSGSDPGICSNSTDSQAOGLKAGKKLKAETLESLDLY 547
Db      1677 NTBLSLTDLKRDITPREG-----STDDTDAKSPMLNPTTVLDEBDKABEEDILA 1725
OY      548 S-----DKKQMSPPICEETDEQAKQLTIDIGSNLEKEJENSRGLE--CRSDPESP1K 599

```

Db 1726 ECHSAMPKSHKRYVKIWDQ-----INHTSAATSSGNSRSMQEBTDKKNKTSVPK 1778  
Qy 600 -----KTSLSPTSLGYSYGRDLDLAKKHAASLQTSDDPADRTTLNH 643  
Db 1779 PMSOSIGKERIKKNTLKNLNPNSNOJCDBPK--PSSKKPSKVAANKITPNNERTYGF 1835  
Qy 644 ADHSKTIYQHR-----LLSROBEL-----KERAVLLEOARRDALKA 681  
Db 1836 AFDS-----PHHTYPIEGTPYCSRYDSSLSLDREDDIDLSEKAKELKEKGTOTDQV 1891  
Qy 682 GNKNNTATATPCNKLSDQODEERRQL--REAROLIAEARSGVKNSLPSYGEAAE 739  
Db 1892 KYKHNRAINPM-----GKODQTPKSLGGRDQKALV-----QKPTSPSSAK 1935  
Qy 740 KUKERSKASGDBENDIEIDTNEBIEPGFVVGGBELITLNDLDTPEONSKLVLDLKKL 799  
Db 1936 GTODRGATDEKMEFAL--ENTPVCF--SRNSSLSL-SDIDQENNNKETEPLK---- 1985  
Qy 800 LEVQPOVANSPPSAOKAVTSSSEODMKSGETEDLATERLOKTERFRAPVVPKSDTVRK 859  
Db 1986 -----QCTSETQLGRPPQTSGYAPKSPHYED-----TPVCFSPRSSLS 2026  
Qy 860 TQLQSPSQYIENREPMKQRSIOEDTKKNEKKAITEORKEPSEDEVLANGFKDTQGY 919  
Db 2027 LSIDSEDDL-----QECISSAMPK-----KRPKS--KIKNEVGKSRMSV 2065  
Qy 920 VGEIALLNEBOKIDTRALYVEKRLRYLMDGRNTEEBEAMQ-----EWMVLNKKAL 974  
Db 2066 GGIIA--EERPLTLDLRDI-----QSPDSNAPSPOSENPDWKAIOGANSI 2110  
Qy 975 IRRMNOISLKEHDLERYELLNRELAMLAIEBMOQTEBAQKREQLLDELVAL--V 1031  
Db 2111 VSRLLQAA-----AAGSISRGSSDSISLSKSGI 2141  
Qy 1032 NKRALVYRDLDAOEKOEEDHELTLEONKG--KMAKKEKCVLQ 1076  
Db 2142 SLGSPFLTLDXEK-----TITSNKGPIILPAKSALE 2176  
  
RESULT 7  
US-10-505-928-357  
; Sequence 357, Application US/10505928  
; Publication No. US20060086532A1  
; GENERAL INFORMATION:  
; APPLICANT: Ludwig Institute for Cancer Research et al.  
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES  
; FILE REFERENCE: 28967/39178  
; CURRENT APPLICATION NUMBER: US/10/505,928  
; CURRENT FILING DATE: 2004-08-27  
; PRIOR APPLICATION NUMBER: US 60/363,019  
; PRIOR FILING DATE: 2002-03-07  
; NUMBER OF SEQ ID NOS: 866  
; SOFTWARE: PatentIn 3.2  
; SEQ ID NO 357  
; LENGTH: 3256  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-505-928-357  
  
Query Match 3.4%; Score 185; DB 6; Length 3256;  
Best Local Similarity 19.1%; Pred. No. 0.0002;  
Matches 231; Conservative 163; Mismatches 470; Indels 344; Gaps 52;  
  
Qy 2 KOYASPMPT---QTDVK-LKFKPLSKVVSALQPSLSCIFLRGKATDEBMOSIASIVS 57  
Db 1172 KAMLTPKRAGDEKDIKAFMGTPVQKLDLAGTLPS-----KROLQTKKKAQLADLAG 1226  
Qy 58 MKQ-----ADICNLDDFEEDNDD-----ENRVNOEKAATIT-ELINKL 97  
Db 1227 FKELFOTPGHTBELVAAGKTKTKIPCDSPQSDPVDPTPTSTKORPKRSIKKADVEGLACR 1286  
Qy 98 NFLDEAKDLATVANSNPDDDAALANPFGDPDSEPTTE--TASPRTEDSFNNNSNP 155

Db 1287 NLMPSAGAMHTPKPSVGEEDIIIF--VGFVQKLDLTENLTGSKRPQ-----TP 1336  
Qy 156 FKELQTPQYLNPPB-----PEAFYTIK-----DSPQC-----STKKNIRPV- 193  
Db 1337 KEKQALIEDLGFKELFOTPGHTBEAVALAGTTMPCSSPSEBADPTPTSTRORPKPTLE 1396  
Qy 194 --DMSKYLVA-----DSG-----KTEEBELDSNFYBPSTYP 225  
Db 1397 KRDVQKLSALKULTQTSGETTHTDKVPGGEDKINAFRETAQKLDPAASVTSKRIK 1456  
Qy 226 PNNLVNPOBL-----ETERRVKRAKAPAPVLSPTKG----- 257  
Db 1457 TKERAOPLLEDLAGKELFOTPVCTDKPTTHEKTKIKACRSQPDVPDPTSSKPSKSLR 1516  
Qy 258 -----VINENTVAGKDLSTSPSPSPSPSYLGRKP--NAOSLILVCKEYTKATYG 307  
Db 1517 KVDVEEBFALRKPTPAGAMHT--PKPA-----VSGBKNIYAFMGTPVQKLDLTENLTG 1570  
Qy 308 VK-----ITNP-----TTSWRGLSPCALIHPRLDLYKSLNP 342  
Db 1571 SKRLQTPKEKAQALIEDLAGKELFOTRGHTBSMTMDTKAKVACXSQPLD---DKNP 1626  
Qy 343 QDIKNNKAYDGPASIGISRLLEPSDMVLAIPLKLTVMTYLYQIRAFPSGOELNV--- 399  
Db 1627 ASSKRLKTS--LGKVGVK-----ELIAY-GKLT-----QTSGETTHHTE 1665  
Qy 400 VOIEENSKTYKGYNTETDNNSSVDQEKYAEISDLKREBELQOPISGAVDPLSQDDSV 459  
Db 1666 PTGDGKSMKAFMBEPKQILDSASLITGSK--ROLRTKSGSEVEDLAGPTE-----L 1716  
Qy 460 FVNSGVGESESEHQ-----PDDHLSPTASPYCART--KSPTEPQ-----KSQ 503  
Db 1717 FQTSHTSEMTNEKTKTVSYRASQPLVDPYTSKQPKSLKADTBESFLAPRKQTP 1776  
Qy 504 SSGRTSGSDPGICSNYDST-----QAQVLGKKRL-----KATLE----- 541  
Db 1777 SAGKAMHTPKRAVEEKDINTFLGTPOVKLDQPNLNGSNRRLQTRKEKAQALBELGFR 1836  
Qy 542 -----LSDLVSDKKKMSBPPTCEBTDEQKQTLDIGNLBEKKEKLENSGLSCRSDE 595  
Db 1837 ELFQTPCTDNPDADEK--TTKKILCKSPQS-----PADPTPTTKORPKSLKAAVEE 1888  
Qy 596 SPIKKTSLP-----TSKLGYYSRD-----IDLAKKHAASLQTSDDPAD 637  
Db 1889 EFLAFKRLTSPAGAMHTPKRAVEEKDINTFVGTPEVKLDLGNLPGSKRRPQTPYEKA 1948  
Qy 638 RT-----TLNHADHS--SKIVORLILSROE-----LKERAVLLEOAR- 674  
Db 1949 KALEDLAGFKELFOTPGHTBSMTDDKITVSCSPQPDVPKPTTSKQRLKISLGKVG 2008  
Qy 675 RDAALKAGKNNTATATPCNROISDQODEERRRLREAROLIAEARSGVKNSLPSYG 734  
Db 2009 KEVLPVQKLTQTSGETTHQTHRETAG--DEKSIKAFKESAKQMDPANYGGMERWPRTP 2066  
Qy 735 EMAEKLKERS--KASDBENDNIEIDTNEBIEPGFVVGGBELTNL-----ENDLTPB 786  
Db 2067 KEKQSLIEDLAGFKELFOTPDHTEESTTD-----DKTKYICKSPPEPSMTPT 2115  
Qy 787 QNSKLVDLKLKLEVOPOVANSPPSAOKAVTSSSEODMKSGETEDLATERLOKTERFR 846  
Db 2116 STRRPPTPYKGDIVE-----ELSAKQLOVTHTTIDKVPGBEDGGINVFRATAQKL 2168  
Qy 847 NPVVFSKDSYVRKTQLQSPSQYIENREPMK--QRSIOEDTKKNEKKAAL----- 895  
Db 2169 DPAAVSYSGKQRPRTPKRGAQPLEDLAGLKELFQTPVCTDKPTTHEKTKIKACRSQPD 2228  
Qy 896 --TETOKPSEDEVLANGFQDQSYVVGELALNEBOKIDTRALYVEKRLRYLMDGRN 953  
Db 2229 VGTPTIFKPSQSKSLRK-----ADVEEELIARKRTPSVGAKMDTPPKAGGD 2275  
Qy 954 TEEBEAM 961

Db 2276 EKDKAKEM 2283

RESULT 8  
US-10-505-928-213  
Sequence 213, Application US/10505928  
Publication No. US20060088532A1  
GENERAL INFORMATION:  
APPLICANT: Ludwig Institute for Cancer Research et al.  
TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES  
FILE REFERENCE: 28967/39178  
CURRENT APPLICATION NUMBER: US/10/505,928  
PRIOR FILING DATE: 2004-08-27  
PRIOR APPLICATION NUMBER: US 60/363,019  
PRIORITY DATE: 2002-03-07  
NUMBER OF SEQ ID NOS: 866  
SOFTWARE: PatentIn 3.2  
SEQ ID NO 213  
LENGTH: 1120  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-505-928-213

Query Match 3.4%; Score 184; DB 6; Length 1120;  
Best Local Similarity 18.9%; Pred. No. 4.9e-05;  
Matches 232; Conservative 186; Mismatches 444; Indels 364; Gaps 56;

Qy 28 AALQFSCTFLRAGKATDEDMQSLASVSKADIGN-----LDDPEENEDDE 78  
Db 47 SSLTPSLCKLGLQGG-----SNSSSPVDFVNNKRTDLSSEHSHSKWLETCCQHSDEQPL 102  
Qy 79 NRVAQEKAKATR-----LINKLNFLEBAKD-LATVNS-----N 113  
Db 103 DPTQISTRTKTSBAVDPLGNVWKTIVLPSPGGQDMIFAKLDTHAETMSISLNG 162  
Qy 114 PF-DDPDAELNP-FGD-----PDSEBITETASPRKTEDSPYNSYNPFKEVQTPOY 164  
Db 163 PLRTDDLVRBEVAVPCMGDRFSEVAVSEKPIFO-----BESPSH 200  
Qy 165 LNPDEBEAVTITKDSPOSTKRNIRPVDMSKTLVADSSKTEBEEL--DESNPFYEPKS 222  
Db 201 L-----LEESPPNCSQ-----LHCSKSLSTRTEAVREEDVPSBSNAFL-PSS 244  
Qy 223 T-----PPN-----NLNVPVQELTERRYKRAKAPAPVLPSPKTGVLENVTSAGKD-L 270  
Db 245 VLMISPESTALADPRVNVHVPDEEIVHGMABEREMAPPTHKSESEEDQALVSSVEDIL 304  
Qy 271 STSPKSPPI-----PSP-----VLGRKPNASOSLLVWCKEVTNRYGVKITNFTTS 316  
Db 305 STCLTPVLVEMESQEGAPVAVEDVGRILGSDTBSGMSPLAWLB-----KGVNTSVMLEN 358  
Qy 317 WRNGLSF-----CALIHRPRLIDYKSLNPDQIKNNKAKADGPAISIG 360  
Db 359 LRQSLSLSPMLRDAIIGTTPSTCSVGWTFPSA-----POKSTNTSQT---GLVG 407  
Qy 361 ISRLBESDWWLAIIPDKLTVMTYLYQIRAHFSGQELNVVQIBNSSKSTYKVGNYETDT 420  
Db 408 TKRSTSTEBQLGCRPDLTALS-----RH-----DLBNL-----L 439  
Qy 421 NSSVDEKFYAEISDLKREBELQOPISGAVDFLSQDDSVFVNDSGVSESEHQTPDDL 480  
Db 440 SSVIVYEFISRQLRDMKSGQLAVPHR-----ETQDSSTQDTDSHGKITNKLQHLKESH- 491  
Qy 481 SPSTASYCRKRTSDTPPOKSGQSSGRTSGSDPGICSTNDSYQAVLLGKKLLAKETL 540  
Db 492 -----EWGALQQA-----RNVVQSWVLISKELI---SL 518  
Qy 541 ELSDLVYSDKKKM-----SPPTICEETD---EOKLQTLIDGSLIBKELNSRSL 569  
Db 519 HLHLHLHEBKTYVNGSSRAETIVCCCFDLKKLRKAKQSL-----KAEBEBAHREMA 574  
Qy 590 CR-SDESPIKKTSLSPTSKLGYVSRL-----DLAKKGHSLRQT 630

Db 575 LRGDAAIVLEAFCAHASQRIQLBQDLASMBEPRGLKDQAQTVGLGHAKQBEIVQOT 634  
Qy 631 ESDPADRTLTINADSSKTVCH-----RLTSROBELKERARVLEQA--RRDAALKXGNK 664  
Db 635 VS-----LTSITLQODWRSQWLDYTTWTALISRRQLTEKLTIVSQQLQGRDAIE--EK 687  
Qy 685 HNTWTATPPCNRQL-----SDQODEERRRQLREBRARQLAARAGVQMSLPSYGEAA 738  
Db 688 QEVSRYLVEQSAQLEBKQGTQTELENTRLADLRQLOIILANMSQLELSQHTHCA 747  
Qy 739 EKLKSKASGDENDNEIDTNEIPEGVVG-----GDELTNLENDLTPE 766  
Db 748 QDLAMDELLCQUTOSNEEQAOQVKEEMALKMQAELQOQQAFLAKEVADLKETLEFAD 807  
Qy 787 QNSKL-----VDLKKLLEVPQVANSPSAQAQAVTESSEQMKSGTELRTERLQ 839  
Db 808 QENQVAHLBVGVECCQKTLTEVLRE--RLQCEMLDYVENTIYAKLASTIANQEBDL 865  
Qy 840 KTERFRANPVPSKDSIVRTQLSFSQYIENPEMKRQRSIQEDTYKNEBRAITETQ 899  
Db 866 KTRQ-----YQKGLGLEQLSLTLPIQTK-----LKEKTEQETLLSTA 906  
Qy 900 RKPSEDEVLNKGFQDTSQYVVELALENQ-----KQIDTR-AALVEKRLRY 946  
Db 907 CPPTQEHPL-----PNRTFFGISLITVADEPESTVPPLIGSDKSAFTRVASVSIQ--- 959  
Qy 947 LMDTGNTBEEBAMQEWFLVNNKNAALIRRMQSLLEKEHDLERYEILNR---ELRA 1003  
Db 960 -----PATPQEMESLLEMSITTELOSL-----CSLLQSK--EPAIRTLQKICELQ 1007  
Qy 1004 MLA:EDMOKTEAQKRE-----QLLL-----DELVALVNRDALVBDLDAQEK--QA 1048  
Db 1008 RLQAQSEQHQVEQKAEADIEKLNQALCLRYKREKELQVYIQNEKILBQIDISGELISL 1067  
Qy 1049 EEEDEHLERTLEONKGMKAKGECVYIQ 1076  
Db 1068 REEVTHLTRSL-----RRAETETK-VLQ 1089

RESULT 9  
US-10-505-928-799  
Sequence 799, Application US/10505928  
Publication No. US20060088532A1  
GENERAL INFORMATION:  
APPLICANT: Ludwig Institute for Cancer Research et al.  
TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES  
FILE REFERENCE: 28967/39178  
CURRENT APPLICATION NUMBER: US/10/505,928  
PRIOR FILING DATE: 2004-08-27  
PRIOR APPLICATION NUMBER: US 60/363,019  
PRIORITY DATE: 2002-03-07  
NUMBER OF SEQ ID NOS: 866  
SOFTWARE: PatentIn 3.2  
SEQ ID NO 799  
LENGTH: 1384  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-505-928-799

Query Match 3.3%; Score 182; DB 6; Length 1384;  
Best Local Similarity 21.1%; Pred. No. 9.1e-05;  
Matches 206; Conservative 115; Mismatches 346; Indels 310; Gaps 42;

Qy 162 PLYLNPDEBEAVTITKDSPOSTKRNIRPVDMSKTLVADSSKTEBEELDESNPFYEPK 221  
Db 161 PPYSPVPAESESIVNGNHTPQATRGPS--ACASHSSLVSSIEKDLQIMSLVLEBEG 217  
Qy 222 STPEPNLVNVPVQELTERRYKRAKAPAPVLPSPKTGVLENVTSAGKD--STSPKSPPI 280  
Db 218 AAGKRAAATSPSPMANGRY-----LISBPT---SPGMSVSSSTVENTSPASPSIS 266  
Qy 281 SPV-----LGRKPNASOSLLVWCKEVTNRYGVKITNFTTSWRNGLSFCALIHHR 330



```

Db 267 SPASSGASHSPGQBPSPVPLVPAR-----SSSYHAL----- 303
Qy 331 RPDLDYKSLNPQDIKENNKAYOGFASIGISRLLEPSD-WLLAIPKLTMTYVOIR 389
Db 304 QPPOSREGASBPRLSRKGHEPPSPGRLGLTDSPPAATVLEAARAPSPRL----- 359
Qy 390 AHFSGOBLVVOIEBNSSKTYKGVNYETDNNSSVDQKFAELSDLKREPELOOPIGA 449
Db 360 ---GGQLPVVAI---SLSEY-----PASGA 378
Qy 450 VDLPLSGDDSVFVNDGVESSSEHQTP-----DHLSPSTASPYCARTSDTEPQ 499
Db 379 ---LSQPTSI-----PSPKFGPPVPAPRNKIGTLQORPPSPFBPBGSRVLT-TSP- 427
Qy 500 KSQOSSGRTSGDDPGICSNSTDSTOAVLGGKRLKAELELSLVYSDKKKMSPPPI 559
Db 428 -SRQLVGTTP-SD--GLATRTLQBPSPRLGRG--DSMBELPLSPSLRRRLS-PLP 480
Qy 560 CEETDEQQLQTLIGSNLEKEKLENSRSLGCRSDPSPK-----KTSLSPTSKLG 610
Db 481 TRITPDK-----LNREVASPRPRMAAHGASBEDSLTIGAGKRTSPSPFLG 531
Qy 611 YVSRRDLDAKKKASLQTESDPDADRTTNHMDH--SSKIVQHRLLSROBELKARAVL 669
Db 532 ESLA-----PHKGSFSGRLSPAYSLGSLTGASPCQSPCYQRLSSGD-----LKVP 577
Qy 670 LEQARDAALKAGKHNNTATPFCNRQ-----LSDQDEERRRQUREAROLIAE--AR 722
Db 578 VTRERKNSITISD--MEDDLLEYHRQORBLRQEMERLERQRLFTILMCAEVSAD 635
Qy 723 SGVMSLEPSYGEMAAEKLKERSKASGDENNIBIDTNEIPEGFVVGSGELTNLENDL 782
Db 636 GGPAGELPSIGETA-----ALALG----- 657
Qy 783 DTPBONSKLVKLKQLEVOPOVANSBSSAQAQVTESSBODKMSGTEJRTERLOKTT 842
Db 658 -----RRSRGLAGASGRSSB-----PGVATQGLWMSB 686
Qy 843 ERFPRPVFSDSTYRKTOLQSPSOYIENRPMKORSIOEDTYKGBEKAALITETOKP 902
Db 687 E-----RDEBNLK-----EBGSSSTESTOGE 707
Qy 903 SEDEVLANKPFQTSQYVVELALENEKOIDTBALVEKRLR-----YLMOTGNTEBEE 958
Db 708 HEDAPSTK-----LOGEVLALBEBDAQVGHVEQIKVAKELBOQLQESABAEHER 759
Qy 959 AMQ-----EWFVNLKKNALIRRMNQLSLE---KEHDLERYEYLINRELRAMLAIED 1009
Db 760 ALLGGERABRALLQKEQKAVDQLOEKLVALETGIQKRDKEA--EALFETKLPEDLE 816
Qy 1010 WQKTBQGRREQ-----LLLD--ELVALVNRDALVRLDLAQ---EKQABEDDEHL 1055
Db 817 FOQERESRVEEBEELAGQGLIRSKABELIRAIKKEKRLALIDSOAQIRAOAVQESRL 876
Qy 1056 ERTLEOKKGNKAKKEK 1072
Db 877 ARDNASLQQLQKEKEK 893

```

```

; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; PRIOR FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2587
; LENGTH: 1674
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-511-937-2587

Query Match 3.3%; Score 180; DB 6; Length 1674;
Best Local Similarity 19.9%; Pred. No. 0.00016;
Matches 206; Conservative 133; Mismatches 322; Indels 376; Gaps 46;

Qy 220 PKSTPPNNLVNPVQBLETERR-----YKKAAPAPVLT 252
Db 53 PTLAPSEPKVDQKLPRSSERRSGCGTQPPARSAVAAGBAARGAAPERGSPLGRRV 112
Qy 253 SPK-----TGLNENTVSAGKDLSTSPSPSPISPVYLGRKNASQ----- 292
Db 113 SPRLCSGEGQAVAVGIAGKRRGRD-GSRAPGRREMFLLHKEFVRQKPPADLRPD 171
Qy 293 SLVWCBEVTKVNGVAVITNFTTSMRGLSP-CALHHPRLDLDYKSLNPQDIKENNK 351
Db 172 BEVYCYKVTNTEIFHYDDFFERTILCNLSWSCAVTG--RP-----GLTYQELSESK 223
Qy 352 AYDFPASI-----GISRLBPSDWLAIIDKLTVMYTYLOIRAFSGQE 396
Db 224 ARQNLQFPPEPLIPVLVYLSLTHSRSLHEICDIPAYVVDORVVEEIVIRANGARLQ 283
Qy 397 LNVQIEBNSSKTYKGVNYETDNNSSVDQKFAELSDLKREPELOOPIGAVPLSQD 456
Db 284 CTILEVLPFPHONGFANGHV--NSVDGE----- 309
Qy 457 DSVFVNDGVESSSEHQTPDHLSPSTASPYCARTSDTEPQKQSSGRTSGDDPGI 516
Db 310 -TITISDS-----DSESTQSCSQNGKKQDALDP-- 337
Qy 517 CSNSTDQAVLGGKRLKAELELSLVSDKKKMSPPICEETDEQQLQTLIGSN 576
Db 338 -----LTFKYK-----VQPTKEHLHSALVKAQISRKRL----- 368
Qy 577 LEKELENSRSLGCRSDPES--PIKTSLS-----PTSKLGYSYSD 616
Db 369 FSRDKLTLFKQHC--SPQGVYIKKASISLYKIAQDPSYFPDDPPFTFSPANRR 426
Qy 617 LDLAKKGHASLRQTESPDADRTTNHADSSTKVQNR--LLSRQELK---EKAIVLE 671
Db 427 GRPEKRIHIS---QEDNVANKQTL--ASYRSKATYKRDKLQKEBKSLAFERAKUKRE 480
Qy 672 QARRDALKAGKHNNTATPFCNRQLSDQODEERRRQUREAROLIAEARSQV--MSB 729
Db 481 KA--DLBAKKK-----EKEDKEK--REBLKVIYEBRLKKEKER 519
Qy 730 LPSYGEMAAEKLERS-----KASGDENNIBIDTNEIPEGFVVG-----GDEL 775
Db 520 LKVEREKERKLEBKRYEYLVKQMSKPRDMECDLKEIPETPVKTRLPEIPFDAL 579
Qy 776 TNE-----NDLTPBQNSKLVDLKXGLLBYQPO 805
Db 580 MVLEFLNAFGLFDLQDEPDPGVTLEVLBALVNSGEP--LCELLFFLTALFOA 634
Qy 806 VANSPPSAQKAVTESSEQ-----DMKSGT--EDLRTE-----RLQK- 840
Db 635 IABEBEVAQQLDADTKGCSLSDLDSCITSEIRLHILASGADVTSANAKYRYQKR 694

```

```

RESULT 10
US-10-511-937-2587
; Sequence 2587, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, Macdonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING

```

QY 841 -----RFRAPVVSXDSYTV-----RKTOLQSPSOYIEN 871  
 Db 695 GGPATDADACMELELNSPISLVKLSSTSVYDLTGECKKILHALCGULTIVSTRDIED 754  
 QY 872 RPEMKRO-----RSIQ-BDTYKKEEKA-----ITETQKPSDEVLN 909  
 Db 755 YVDILRAQKOFRELEKAKOHRKEREBAARIRKREKEKQEQKMEKOEKLEDBORN 814  
 QY 910 KGFPTDSQYVVGELALBNQKQIDTBAALVERKRLRLMDPTGRTEEBEAMQEWPLVN 969  
 Db 815 -----STADISIG-----EEREDDTISIESKDTQKEL-DQDMTEDEDDPGS-----H 858  
 QY 970 KXNALIRRNQOLSLIEKHDLERRRYELNRELAMALIEDWQTEAQRREQLLDEL-- 1027  
 Db 859 KRG-----RGRKGQNGFPEFTROBOINCVRHEL--LTDBEALAKQEHQREKELLEKIOS 913  
 QY 1028 -VALVN-----KRDALVR 1039  
 Db 914 AIACTNIFPLGRDMYR 930

RESULT 11  
 US-11-118-524-2  
 ; Sequence 2, Application US/11118524  
 ; Publication No. US20060088847A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CU WEI  
 ; TITLE OF INVENTION: ARF-BP1 AS MEDIATOR OF P53-DEPENDENT AND INDEPENDENT TUMOR  
 ; FILE REFERENCE: 19240-497US2  
 ; CURRENT APPLICATION NUMBER: US/11/118,524  
 ; CURRENT FILING DATE: 2005-04-29  
 ; PRIOR APPLICATION NUMBER: 60/610,506  
 ; PRIOR FILING DATE: 2004-09-15  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: PatentIn version 3.3  
 ; SEQ ID NO 2  
 ; LENGTH: 4373  
 ; TYPE: PRT  
 ; ORGANISM: homo sapiens  
 US-11-118-524-2

Query Match 3.0%; Score 167; DB 7; Length 4373;  
 Best Local Similarity 18.2%; Pred. No. 0.005;  
 Matches 206; Conservative 164; Mismatches 450; Indels 310; Gaps 41;

QY 45 TDEDMOSLASVSMKQADIGNLDDFEEDNEDDENRVQSEKAKITELINKLFLD-EA 103  
 Db 2275 SEODAQASQDSSSNODPGEBAEVOEEDHD---VTQTEVADG-----DIMGEA 2323  
 QY 104 EKDLATVNSNP-----PDDPAALNPNFGDDSEBPIETASPKRTBSFY 149  
 Db 2324 ETYSVVIAGQPEVLSQEMQVENLELDLLELRRDQSGSNSTIIVRSGEBSQEBVLN 2383  
 QY 150 NNSVNPFEKVOQTPQ-----YLPPEDEPAFTYIKDSPQSTKRNKINPVMSKYLVA 201  
 Db 2384 DEAPSVLSQASTIQAKREDSEMLIDPEDEEH--TQSEBDSGSGNEEDD----- 2429  
 QY 202 DS-SKTEEBEELDSNPFEKSTPPPNLVNVOELETERRVYRKAPAPVLSPKTG--- 257  
 Db 2430 DSDQEEEBEEDDEDEDEGEDEDDDDGSEMELDDYDPMNASPLVRFERFRED 2489  
 QY 258 ---VLENVYSAKDLSTSKSPISBPVYGRKPNASQSLVWCKEYTKYRGVKITNFT 314  
 Db 2490 DLIIEFDNMFSSATDIPPSGNIPTHPLWVRHADHSLTIGSGSSSTTRLTQIGISQRT 2549  
 QY 315 TSWRNGLSFCALIMHF-----RPLLDYKSLNPQ---DIKENNKAYADGPAISGTSRL 365  
 Db 2550 LRQULTANTGHTIVHTVGNRQPNPPLILQRLLEPSAADLLQSSSL--PLQSRGRAKEL 2607  
 QY 366 EPESDMV-----LALIPDKLTWVTYLYQIRAHFSGQELN 398

Db 2608 VGNDVHIIARSDELDDFFHDQSTATSQAGTILSIPALTRWT-----BECK 2656  
 QY 399 VQOIEE-NSSKSYKYKYNVETDINSSVDOKEFYAEISDLKREBELOQPIGAVDFLSQDD 457  
 Db 2657 VLDAESMHDVSVVKV-----STVNHLEFL-----RDELEERREKRRKQALAE 2701  
 QY 458 SVEFVNDGCVSESESEHQTPDDHLSPTASPYCRRYTSDETPQSQSGRTSGSDPGIC 517  
 Db 2702 TK-ITDKGDEKKNRDQ-----SAGCTASNSNDTBQN-LSDGTFMDSYTTTP 2748  
 QY 518 SNTDSTQAOVLLAKKRLKAEYELSDLYVSDKKOM-SPPFICEETDEQKLTLDIGSN 576  
 Db 2749 SSTDAATSE-----SKET--LGLTQSSQOQPTLPTPPALGAEVQOELOSAGGEGSS 2797  
 QY 577 LEXEKLENSLSLECRSDPSPPIKKTSLSPSKIGYVSRDLAKKHAASLRQTESPDA 636  
 Db 2798 TOL-----LMPVEBELGPFRGEBALFTQWELSPATTISLSPERAEDS 2842  
 QY 637 DRTTLNADHSSKIVOHRLLSROEELKERARVLLQARPDALKAGNKHTNTATPFCNR 696  
 Db 2843 DALT-----AVSSQLEGSFMDTSSLASCTLEBAVGDTS-AAGSS- 2881  
 QY 697 QLSDDQDEERRROLRBRARQLIAEASGVKMSLPYGENAAEKERSVAG----- 749  
 Db 2882 -----QPRAG--SSTPDADAPPAVAEVOGRSDSGESAQPR 2915  
 QY 750 DENINIEIDTNEIEIPGFVVGGEDELTNLEND--DTPENGSLTVOLK----- 795  
 Db 2916 DSSPASPSSSTRDSAVALSAGDSRGILBEPSTSSBEBDLAIGSLPEGVDPFPLAA 2975  
 QY 796 -----LKKLLEVOQVYANSPSS-----AAQAVTESSEQDMKSGTBDLATERL- 838  
 Db 2976 LPDDIRREVLYQNGLRIPRTKAPSTNSSAPAVNGPCTEVSEPELALPRAIQSEVLA 3035  
 QY 839 QKTERFR-----NEVFE--SKDSTVKTQLOSFSQYI----- 869  
 Db 3036 QORAEQORRELAQNASDPTMDEVTFLQTLPSDLRSVLEDMEDSVLAAMPPIAABAQ 3095  
 QY 870 ENRPEMKRQROSTOBDTKKNEEKAITETQKSEBEVL--AKGFDQTSQYV-----VGE 923  
 Db 3096 LRREQARQRLMHEHLFGHSITSALSLRSAPFTSRLSGNGVGYTRLAVGRGTFQM 3155  
 QY 924 AALENEQKQIDTBAALVERKRLRYLMD-----TGRNTEEBEAMQEWFMV 968  
 Db 3156 GSSSSNNRPSGSDTILRLGRILLDHEALSCVLVLFVDEPLANTSLRLVRLNCYHA 3215  
 QY 969 NKXNALIRRNQOLSLIEKHDLERRRYELNRELAMALIEDWQTEAQR 1018  
 Db 3216 QTRHWVIRSL--LSIIQSSSESE-----LCIETPPLTTSSEK 3250

RESULT 12  
 US-10-511-937-2597  
 ; Sequence 2597, Application US/10511937  
 ; Publication No. US20060088836A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: EXPRESSION DIAGNOSTICS, INC.  
 ; APPLICANT: Wohlgemuth, Jay  
 ; APPLICANT: Fry, Kirk  
 ; APPLICANT: Woodward, Robert  
 ; APPLICANT: Ly, Ngoc  
 ; APPLICANT: Prentice, James  
 ; APPLICANT: Morris, MacDonald  
 ; APPLICANT: Rosenberg, Steven  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING  
 ; FILE REFERENCE: 506612000104  
 ; CURRENT APPLICATION NUMBER: US/10/511,937  
 ; CURRENT FILING DATE: 2004-10-19  
 ; PRIOR APPLICATION NUMBER: PCT/US2003/012946  
 ; PRIOR FILING DATE: 2003-04-24  
 ; PRIOR APPLICATION NUMBER: US 10/131,831  
 ; PRIOR FILING DATE: 2002-04-24

PRIOR APPLICATION NUMBER: US 10/325,899  
 PRIOR FILING DATE: 2002-12-20  
 NUMBER OF SEQ ID NOS: 3117  
 SOFTWARE: PatentIn version 3.2  
 SEQ ID NO 2597  
 LENGTH: 851  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-937-2597

Query Match 3.0%; Score 166; DB 6; Length 851;  
 Best Local Similarity 17.8%; Pred. No. 0.00052;  
 Matches 148; Conservative 146; Mismatches 284; Indels 252; Gaps 34;

QY 403 EENSSEKSYKYGNTVDNSSVDOEKFAELSDLKREBELQOPISGANDPISQDSDVAVN 462  
 DB 84 BDSKALSOXKJIKKKTQYKEIDSNVLEKKETK-----AN 121  
 QY 463 DSGVSESESEHQT- DDLSPSTASPYCR-RTKSDTEPOKQSSGRTSGSDDPICANT 520  
 DB 122 HKG-----SENKTDUNSIGIKMKTSCKFKIDSNISKDKSKERTQNKKEKKNIVQHT 176  
 QY 521 DSTQAQVILGKKRLKATLELSDLVSDKKKMSPPICEYDE-----OKLTLDI- 573  
 DB 177 TDSS-----LEKQRTLDSTGSEIV-----KSPRIECSKTRREMOSVQVLMTRDSD 223  
 QY 574 -----GSNLEKEL-ENSRSL-ECRSDESPSPIKTSLSFTSKLGYSDLDLAKKH 624  
 DB 224 GYENSTDCMCKALBERSESSEISGDSSENEITSVGRA-----SDDDSDSEDE 276  
 QY 625 ASLRQTESDPADRTTLNHAHSSKIYQHR--LLSROBELKERRAVILLEQARRDALAKG 682  
 DB 277 EEDDEDEDEDEDEDDKSDSGPLANGKNIEISSDEDDTDADLPFBESGFEHAMEL 336  
 QY 683 NKNNTATPFCNQLSDQODEERRRL-----REBARQLIA-----EARSQVMSL 730  
 DB 337 DKD-----APRADEITRLAVCNMDWRLAKDLALFNSPKFGVTFSVK 383  
 QY 731 PSYEMAAEKLEKERSKASGDENDNIEIDTNEIPEGPVVGSGDELTYNEN----- 780  
 DB 384 IYSEFGERKEKEOVQVPELISL-----EDAPE-----KQWTSREKLDYQFKLK 432  
 QY 781 -----DLDTPEONSKL-----VDLKL-----KKLLEVQPVAN 808  
 DB 433 YTYAVVDCDSPTASKIYEDDCGLFESSCSFIDIRFPDITDPDEKDVASRYNLTAY 492  
 QY 809 SPSSAQAAYTESSEBODKSGTEBILRTERLOKTERFENPVVFSKSTVRKTOLQSFQY 868  
 DB 493 KPKFTSAAGTSTVBIITWEDTH---ERITMLNRKFKKEBLMD-----PGAY 539  
 QY 869 I-----ENRPEMKRQ-----SIQDITK-----KNEBKAA-----ITETQKPSD 905  
 DB 540 LASSSEDEBEIEBELQCGDGVNVEDGKTKKQODEBQIAKRYQLQVIOKEKEKREN 599  
 QY 906 EVLNR-----GFKDTQVVGELALENEKQIDTRALVYERLRYLMDTRNTBEERAN 960  
 DB 600 DMKEIKTVPELKESASBEVKN--KLBSGDKLTPWEOFLKKEKKEKRLKROKALAEAS 657  
 QY 961 MQE-----WFL-----MLVKNKALIRMANQSLLEKEHDLERREYELNRELFRAM 1004  
 DB 658 EEBELPSVDLNDLPFAEBVQIGINKS--VKSADGSPBEEIEIERQ-----KAEWMLL 711  
 QY 1005 LAIED-----WQTEBAQRRBQLLDELVALVN----- 1032  
 DB 712 MMDEDEBSKHFNNYKIVEHQNTLSKKKKQOLMKKKELEDFEVAVVNDARFOAMYSHLF 771  
 QY 1033 -----KRDALVRDLDAQBEDEBHLBERTLEONKNGKAKKEK 1072  
 DB 772 NLDSDFPKFKTKYMEKILBEKAROREKEBELTOAIKKESELEKESQ 821

RESULT 13  
 US-10-196-749-346

Sequence 346, Application US/10196749  
 Publication No. US20060094864A1

# GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
 APPLICANT: Chen, Jiah  
 APPLICANT: Deanoyers, Luc  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Goddard, Paul J.  
 APPLICANT: Gunney, Austin L.  
 APPLICANT: Pan, James  
 APPLICANT: Smith, Victoria  
 APPLICANT: Watanabe, Colin K.  
 APPLICANT: Wood, William I.  
 APPLICANT: Zhang, Zhen

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ACIDS ENCODING THE SAME

FILE REFERENCE: P3430R1C340

CURRENT FILING DATE: 2002-07-16

PRIOR APPLICATION NUMBER: 10/052586

PRIOR FILING DATE: 2002-01-15

PRIOR APPLICATION NUMBER: 60/059263

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/059266

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/063120

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063121

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063486

PRIOR FILING DATE: 1997-10-21

PRIOR APPLICATION NUMBER: 60/063540

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063541

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 612

SEQ ID NO 346

LENGTH: 671

TYPE: PRT

ORGANISM: Homo Sapien

US-10-196-749-346

Query Match 3.0%; Score 164.5; DB 6; Length 671;

Best Local Similarity 22.0%; Pred. No. 0.00047;

Matches 145; Conservative 102; Mismatches 260; Indels 151; Gaps 31;

QY 465 GVGSESEHQTDDHLSPTASPYCRRKSPTEPOKQSSGRTSGSDP--GIC----- 517  
 DB 76 GFNEGLWEIQW-NPHAYSAAPPV---SSDSSEAPBANPADGDADEDDRGVAVATVAV 131  
 QY 518 -----SNTDSTQAQVILGKKRLKATLELSDLVSDKKKMSPPICEYDE 567  
 DB 132 TATAASDRMESDSDSDSNGLKR--KTPALMS--VSKARKAS-----SDLQOAS 181  
 QY 568 LQTDIGSNIEKLENSRSLSECRDPE-SPIKTSLSPTSK--LGYSDRLDLAKKH 624  
 DB 182 V-----SPSEBESSESEKTSDDPTPEKKAARARPRGLG-----GRKKKA 228  
 QY 625 ASLRQTESDPAD-----RTTLNHAHSSKIYQHRLLSROBELKERRAVILLEQAR 674  
 DB 229 PSADSDSKADSDAKPEPVAMARASASSSSSSSDSDVSVKPPRGRKPAEKPLPKPR 288  
 QY 675 RDAALTKGNNTATPFCNQLSD-----QDEBRRQLRERARQLIAEAS 723  
 DB 289 -----GRPKPE-RPPSSSSSDSDSDVDRISSEWRDRARRRREARRRQEE--- 337  
 QY 724 GVKMSLPSYGEAAEKLERSKASGDENDNIEIDTNEIPEGPVVGSGDELTYNLENDLD 783

Db 338 -----ELRLREDEKEE-KERRRERADRG-----EAERSGSGSSGDEL-EDDER 380  
Qy 784 TPEONSKLYVLDLKLKLELVOP--QVANSPPSAQKAVTESSEODMKSGTEDLRTERLOKT 841  
Db 381 VKRGRGRGRGPRSSSDSEPEALEREAKKSAKPPSSSTEPARKGQCKEVRPEKQ 440  
Qy 842 TERPRNFVPSKOSTYKTKLOQSFQYIENRPEMKQRSIOEDTKKNEE-KAAITETOR 900  
Db 441 Q-----AKPVKVRTRKRSEGFMDRKEVCKEKSVEKLOKHSEIKFALKVD-- 489  
Qy 901 KPSEDEVANKGFOTSGYVVGELALENE---OKQDTRALVYKRLRYLMDGRTTER 956  
Db 490 SPVYKCLN-----ALBEIGTLQVTSOILQKNTDVATL-KKLRRYKANDVNEKA 539  
Qy 957 EEAAMQEFMLVKNKALIRBMNQLSLEKEDLERERYELLNRELAMLAIEDMOQTEA- 1015  
Db 540 AEVYTRAKSKVLPKITAQVKNKAG-MEKEKAERK-----LAGEBELAGEAP 586  
Qy 1016 QKREQLLDLDELVALVNRKDALVRDLDAQEKQAEDEDEHLERTLEONKGMKAKKEK 1073  
Db 587 QKRAEDKPTDLSAPVY-----GEATSGQGESAEDEKHEE-----GRDSEEGPRC 631

RESULT 14  
US-11-252-276-22  
Sequence 22, Application US/11252276  
Publication No. US20060094868A1  
GENERAL INFORMATION:  
APPLICANT: Giuliano, Kenneth A.  
APPLICANT: Bright, Gary  
APPLICANT: Burroughs-Tencza, Sarah  
TITLE OF INVENTION: A System for Cell Based Screening  
FILE REFERENCE: 97-022-K2-CO  
CURRENT APPLICATION NUMBER: US/11/252.276  
PRIORITY FILING DATE: 2005-10-17  
PRIORITY FILING DATE: 09/713, 572  
PRIORITY FILING DATE: 2000-11-15  
PRIORITY FILING DATE: 09/430, 656  
PRIORITY FILING DATE: 1999-10-29  
PRIORITY FILING DATE: 09/398, 965  
PRIORITY FILING DATE: 1999-09-17  
PRIORITY FILING DATE: 1998-02-27  
PRIORITY FILING DATE: 1998-02-27  
PRIORITY FILING DATE: 1997-02-27  
PRIORITY FILING DATE: 1997-02-27  
PRIORITY FILING DATE: 1999-05-26  
PRIORITY FILING DATE: 60/106, 308  
PRIORITY FILING DATE: 1998-10-30  
NUMBER OF SEQ ID NOS: 168  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 22  
LENGTH: 1610  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:  
US-11-252-276-22

Query Match 3.0%; Score 164; DB 7; Length 1610;  
Best Local Similarity 18.3%; Pred. No. 0.0018;  
Matches 186; Conservative 127; Mismatches 404; Indels 298; Gaps 37;

Qy 4 VASPMFTQDVVKIKFPLSKKVSAALQPSLSCTFLREGKATNDEMOSLASLVSMKQADI 63  
Db 204 YOSALSKDNEKNDHNVLLFVTAAGITLGMDELY-KKGDVVD-----GMADL 250  
Qy 64 GNLDDDEEDNEDDERVNOEKAAK-----ITELINKNFLDEAKQDLATVNSNP 115  
Db 251 SLVDALTEP-PPETGEIKRDFMAALBAEPYDIDVGEIVTEKTEFPLDDEKTSSEK 309

Qy 116 DDP--DAEALNPFQDPS-----EPITETASP-----RKTEDSFYNNSYNPFK 157  
Db 310 KRCCLDTSQVE--GIPSSKPTLLANGDHGEGNNTAGSPFDLEBRVDYDQSSQWPE 367  
Qy 158 EVQ----TPQYLNPFDEEAVTTIKDSPPOSTKKNIRPVDMSKLYADSSKTEBEELDES 214  
Db 368 DASFCFQPOQVLDTDQAEFPNEHRDDGLDLLFVSSGPTVASAFTERDNPSEDSYGMFLPC 427  
Qy 215 NPFYE-----PKSTPPNNLVNPFQLETERVRVKAAPAPVLSKPTGV----- 258  
Db 428 DSPASTAVUSQEWVAGAPNSPCSESCVSPETVETTLQPATLSKAAEVSVKQLPAKAL 487  
Qy 259 --LNENTVSAKDLSTSPKPSPIPSVYGRKPNASOSLWVCKEVTKNYKXITNFTTS 316  
Db 488 ETMAEQDTDVVHSPSTDTTGPDTBAALKD-----IDEIK----- 524  
Qy 317 WRNGLSPCALHHPRLDLDYKSLNPODIKENNKAAYDGPASIGISRLPE---SDMVL 373  
Db 525 -----PDVI-----LANVTQPTSESDMFLA 544  
Qy 374 AIPDKLTWMTYLVQIRAFSGOEL---NVYQIEENSXSTYKV-----GN----- 415  
Db 545 QDMELLT-----GTEAIAHANNITLPTBEDSSTKQVADPMESEIVPGNDTTS 591  
Qy 416 -YETDNTSSVDQEKFYAEISDLKREPELOPISGAVDFLSQDDSVFVNSGVGE----- 468  
Db 592 PKTEETTLPIKMDLAPBEDVLTIKETEL-APAGWVSLSIEBALAKNDVRSALPIVAQE 650  
Qy 469 ---SESEHQTPDDHLSPSTNAPYCRRTKSDTEBQKQOQSSGRTSGSDDPICNTDSTQA 525  
Db 651 TVVSETEVVLATVAVLPS--DPYTLTKOVTLPLBAER-----PLVTMTSTLET 698  
Qy 526 QVILGKKRLKAEATLEISDLVYSDKKKMSPPICET-----DEOQLQTLIDGSNL 577  
Db 699 EMTLQKETAPEYET---NLGMA--KDMSELPSEVTLAKQVYILPETKVAEPNVTET 751  
Qy 578 EKEIKENSRSLBCRSDPESPICK-----TSLSPYSKLGVSYSRDLDAKKK 624  
Db 752 SEEHVTSVDMQSPSAETAEPALAKNADLHSGTELVIVNSMAPADLALPLETKVATVPYKD 811  
Qy 625 ASLRQTESPPDADRTTLNHAHSSKIVQHRLLSROBELKERARVLLQARRDALQAKNK 664  
Db 812 KGTVOETBEKP-----REDSQL-ASMQ 831  
Qy 685 HMTNTATPCPNRLSDQOQDEBRRRLQERARQULIAEASGVKMSLPSYGEMAAEKLER 744  
Db 832 HKQGSTVPPTCTASPP-----PVKAAEQMSTLPIIDAPSPLENLBOK 871  
Qy 745 SKASGDENDNIE-IDTNEBIPBGFVVGGDDELTNLENDDTP--EQNSKLVLDLKLKLE 801  
Db 872 ETPGSGPSEPCSGVSRQBEAKAAVGVTG-----NDITTPNKKEPPSPBKKAPLAT 923  
Qy 802 VQPVANSPPSAQKAVTSSSEODMKSGTEDLRTERLOKTTTERFNPNVYFSKOSTYKTKO 861  
Db 924 TOP--AKTSTSAKXCTPISLPKQAPPTTSGLN-----KKPMSLASGVPAAAP 970  
Qy 862 LQFSQYIENRPEMKQRSIOEDTKKNEKAITETQ-RKQSEDEVLVKGFOT 915  
Db 971 KPPAATATATPSTLPARDVK--PKPTIAKVAEKRTSPSKSSAPALKPGPKTT 1023

RESULT 15  
US-10-511-937-2566  
Sequence 2566, Application US/10511937  
Publication No. US20060088836A1  
GENERAL INFORMATION:  
APPLICANT: EXPRESSION DIAGNOSTICS, INC.  
APPLICANT: Wohlgenuth, Jay  
APPLICANT: Fry, Kirk  
APPLICANT: Woodward, Robert  
APPLICANT: Ly, Ngoc  
APPLICANT: Prentice, James  
APPLICANT: Morris, MacDonald

APPLICANT: Rosenberg, Steven  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING  
 TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION  
 FILE REFERENCE: 506612000104  
 CURRENT APPLICATION NUMBER: US/10/511,937  
 PRIOR FILING DATE: 2004-10-19  
 PRIOR APPLICATION NUMBER: PCT/US2003/012946  
 PRIOR FILING DATE: 2003-04-24  
 PRIOR APPLICATION NUMBER: US 10/131,831  
 PRIOR FILING DATE: 2002-04-24  
 PRIOR APPLICATION NUMBER: US 10/325,899  
 PRIOR FILING DATE: 2002-12-20  
 NUMBER OF SEQ ID NOS: 3117  
 SOFTWARE: PatentIn version 3.2  
 SEQ ID NO 2566  
 LENGTH: 1186  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-511-937-2566

Query Match 3.0%; Score 162.5; DB 6; Length 1186;  
 Best Local Similarity 18.4%; Pred. No. 0.0015;  
 Matches 221; Conservative 167; Mismatches 417; Indels 395; Gaps 53;

QY 2 KOVASPMPTOTDVKLKFKPLSKKVVSALOPSLCIFIREGKATDEDMQSLASIVSMKQA 61  
 DB 95 KOLASSDSRKTTEKLTFLKQAIKTAFR-----SKRDEALPSLTQV--RREN 141  
 QY 62 DIGNLDPEED---NEDDDE---NRVQEKAKITELINKNLFDEAEKDLATVNSN 113  
 DB 142 DLYVLPLOGBEKKSSSEEDKEWGERMNO--KQALQSEFFHNPAIDISEDESSL--- 196  
 QY 114 PFDDPDAAELNPFQDPSEBPTF-----TASPKTEB----- 146  
 DB 197 ---PPEV-----KHRIITDKEFTKRRRTLFKAMPESSDPFGYQKGLKKNYL 243  
 QY 147 -----SFYNNSYNPFKEVOTPOYLNPFDEBEAFVTKDSPPOSTK 186  
 DB 244 NQHIENHVKENQOOSHIRQYEDGEGFLKESRRVVS--EDTSYIILKIGIOAKTVA 301  
 QY 187 RKNIRPVMSKLYVADSSKTEEBELDESNNPYEPKSTPP--PNNLV-----NPVQEL 236  
 DB 302 EVDSESLPSSSKMGMSPDVKSSPCBKLTKEBPDPATPPSPRTLLAQALLGSSSEEL 361  
 QY 237 ETERKRV--RKAP-----PPVLSPTKGLVLEN--TVSAGKULSTSPKSPPIPS 282  
 DB 362 ESENRRQARGNAPAAVDEGISISPTLSAIRKALDDEDVKVCAGDDVQGT----- 412  
 QY 283 VLGKRPNASOSLLWCKEVTKNY--RGVKITNFTTSMRGLSPCALIHHFRPDLIDYKSLN 341  
 DB 413 ---GPGAEKRT--NSSTENSDEGLKVRD---GKGIPTPTATLAS-----SSVN 452  
 QY 342 PQDIKENNKAYDGPASIGISRLLEPSDMVLAIIPDKLTWTVYLYQIRAHFSGOELNVQ 401  
 DB 453 -----SAEHNVASTNEGR--EPIDSV-----PKGQMSLVH 480  
 QY 402 IENSSKSTYVGVNVTDTNSSVDOEKYALSDIKREPELQOPIGAVDFLSODDSYFV 461  
 DB 481 -----VGTFAF--PISD-----ESMIKDRKRLPLESAVVR 509  
 QY 462 NDSGVGSESEHQTDDHLSPTASPYCRKTKDTEPOKSOSSGRTSGSDDPGICSNTD 521  
 DB 510 HSDAPGLNGRELTP-----ASPTCTNSVSKNE----- 537  
 QY 522 STQAVLLGKRLKKAETLALSDLYVSDKKOMSPPICEETDROKLTQTLIDIGSLKRGK 581  
 DB 538 -THAEVLEQONELCYEBSKPDSSILSDDETKCKENASAVIIGPVSLQETSIYSVPSBA 596  
 QY 582 LENSRSLECRSDPSPPIKKTSLSPTSKLGYSYRDLIAKKHNSLQTESDPDADRITL 641  
 DB 597 VDNVENV-----VSFNAKEHENFLETTIOEQQTESAG 628  
 QY 642 NHAHSSKIVQHRLLSROBELKERARVLLBOARRDAALKAGNKNTNTATPFCNRQLSDQ 701

DB 629 QDLISIPKAVPMEIDSESDSDSPFIEQSVISDEELQA-----EPFETSKPSEQ 680  
 QY 702 QDEERRKQERARAQOLIAEASGYMSLPSYGYMAAEKLEKRSKASGDEDNIEIDPNE 761  
 DB 681 GEER-----LV-----GTREGAPAESSESLRDNSBRDVEDG--EPOEAKQAD 723  
 QY 762 EIPGPFVVGGDDELTNLENDLDTPEONSKYVDLKLKK-----LLEVO----- 803  
 DB 724 SLHEWQDI--NLEELFTLESNL--LAQONS---LQAQKQOQERIAATVYGMPLESQELLR 777  
 QY 804 ---POVANSPPSAQAQAVTESSEODKMSGTEDLRTERLQKTERFRNPVVSQDSTYRK 859  
 DB 778 LFGIPYIQAPWEBAQAQALIDLTDQTSGETTDD--SDIWLFGARHYNR--FENKXKFV-- 832  
 QY 860 TQLOSPSOYLE--NRPEKQKRSIQ-----EDTKKNEERQAALTETQKRSBDEVLANK- 910  
 DB 833 ---EYQYVDFFHQGLGDRMKLJNLAVYLGSDYTBGIPYGVCTYAM-----ELNBEF 881  
 QY 911 ---GFKDTSOVVVGELALLENBOKQIDTRALVYERKLRVYLDTG--RNTSEBAMQOEWF 965  
 DB 882 PCHGLEPLKFSWMWHBQKNPKIRPNBDTKVKKLTQTLTGFFNPVAVAEALYKP-- 939  
 QY 966 MLVNNKKNALIRNNQLSLEKEHDLERRRYBELNRELPAMLAIEDWQKTEAQKREBQLILD 1025  
 DB 940 -VVDSDSKG-----SFLWGPDLDDKIRFCQRYF-----GNNRT--KTDESLEF-- 978  
 QY 1026 ELVALVNRDALVDDLAQEKQ-----AEEDEHLEFTLRONKQK---NAKKEK 1072  
 DB 979 -----FVLKQLDAQQTQLRIDSFRLAQQEKEDAKRIKSQRLNRAVTQMLRKKEK 1028

Search completed: May 20, 2006, 17:30:35  
 Job time : 18 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2006, 17:09:27 ; Search time 49 Seconds  
(without alignments)  
2112.842 Million cell updates/sec

Title: US-09-856-723a-8

Perfect score: 5489

Sequence: 1 MKQYASPMPTQTDVAKFKP.....RTLEQNKGMKAKKEKCVLQ 1076

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 80: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	402.5	7.3	901	2 T20122	hypothetical prote
2	354.5	6.5	1738	2 T14867	interleukin - slime
3	326.5	5.9	3433	1 S28381	utrophin - human
4	324.5	5.9	3488	2 T34418	hypothetical prote
5	321	5.8	862	1 PADOAA	alpha-actinin - sl
6	319	5.8	2128	2 I52577	beta-spectrin - mo
7	318.5	5.8	1790	2 S67593	transport protein
8	316.5	5.8	2291	1 A46147	spectrin beta chain
9	316.5	5.8	4063	2 T42993	probable spectrin
10	316.5	5.8	4101	2 T23630	hypothetical prote
11	315.5	5.7	2137	1 S4HUB	spectrin beta chain
12	309	5.6	2326	2 T29140	hypothetical prote
13	299	5.4	2139	2 T18296	myosin heavy chain
14	297.5	5.4	3678	2 S28916	dystrophin - mouse
15	297	5.4	3187	2 JC5837	364K Golgi complex
16	292	5.3	2388	2 JE0271	beta spectrin, bet
17	289.5	5.3	2364	1 A44159	spectrin beta-G ch
18	286	5.2	2116	2 A26655	myosin heavy chain
19	285.5	5.2	3660	1 S02041	dystrophin, muscle
20	285	5.2	894	2 T26149	hypothetical prote
21	280	5.1	894	2 T13413	alpha-actinin 2 -
22	279	5.1	924	2 T13413	probable alpha-act
23	279	5.1	1645	2 A37792	spectrin beta-H ch
24	278	5.1	1909	2 A45592	liver stage antigen
25	275.5	5.0	895	2 T13414	probable alpha-act
26	273	5.0	895	2 T13414	alpha-actinin - fr
27	273	5.0	992	2 T46337	hypothetical prote
28	272.5	5.0	1818	1 S73852	hypothetical prote
29	271	4.9	897	2 S02032	alpha-actinin 2, s

30	270.5	4.9	2094	2 S33124	tpi protein - huma
31	268	4.9	1526	2 A45605	mature-parasite-in
32	267.5	4.9	854	2 T17288	hypothetical prote
33	267	4.9	1034	2 T32297	hypothetical prote
34	267	4.9	3685	1 A27605	dystrophin, muscle
35	266	4.8	911	2 JC7186	alpha-actinin-4 -
36	265	4.8	5327	2 T13564	microtubule-associ
37	263	4.8	2057	2 S61477	myosin II heavy ch
38	262.5	4.8	2663	1 S28261	centromere protein
39	262	4.8	1133	2 T22976	hypothetical prote
40	260	4.7	1164	2 T24806	hypothetical prote
41	259	4.7	1927	2 G64585	cag pathogenicity
42	258.5	4.7	904	2 S45673	alpha-actinin, 115
43	258.5	4.7	1093	2 A47212	transcription fact
44	258	4.7	1939	2 T18372	repeat organellar
45	257.5	4.7	2396	2 T13714	kakapo gene protei

#### ALIGNMENTS

RESULT 1  
T20122  
hypothetical protein F25B3.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T20122; T21324  
R:McMurray, A.  
submitted to the EMBL Data Library, April 1996  
A:Reference number: Z19225  
A:Accession: T20122  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-901 <WT>  
A:Cross-references: UNIPROT:Q18749; UNIPARC:UPI00000772F3; EMBL:Z70750; PIDN:CAA94750.1;  
A:Experimental source: clone CS0F4  
R:Gardner, A.  
submitted to the EMBL Data Library, April 1996  
A:Reference number: Z19406  
A:Accession: T21324  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-901 <WT>  
A:Cross-references: UNIPARC:UPI00000772F3; EMBL:Z70752; PIDN:CAA94758.1; GSPDB:GNO0023;  
A:Experimental source: clone F25B3  
C:GeneID:8  
A:Gene: CESP:F25B3.1  
A:Map position: 5  
A:Introns: 25/3; 47/2; 199/1; 242/1; 315/2; 544/3; 600/1; 626/1; 646/1; 666/3; 706/3; 75

Query Match  
Beat local similarity 20.1%; Pred. No. 7.3e-10;  
Matches 219; Conservative 159; Mismatches 330; Indels 381; Gaps 43;

QY	11	QTDVKKFKPLSKVNAALQFSLSCFLRGGKATDEDMQSLASLVSMKADTGNLDFE	70
DB	132	KSEIKLGRPTPLKLCSLIILGSSLSIRVNAADVITSSAASVPEKPVTVDEVAS	191
QY	71	EDNEDDENRVOBEKAKITELINKLPDDEAKOVLATVSNPFDDPDAELNPFDDP	130
DB	192	ANDRRAAAQQIDTITTKMWEIKE-----EKETA-----PRVFA	228
QY	131	SEEPITETASPRKTEDSFYNNYNPKFVOTPOLNPFDEEAFVTKDSPPOSTKKNKI	190
DB	229	HRKSGENENPAKVR-----PQWRVAEER-----KTSQPAA-----	260
QY	191	RPVVMKTYIADSKTEBELDSNPFYEPKSTPPPNVYNPVQELTEHRRVKAKAPAP	250
DB	261	PADIATNRSRSRRASTTEKASTTIFRVDN-----SHHVEVTA	300
QY	251	VLSPKTVLNNETVYSAQKDLSTSPKPSPIPSVPLGRKPNASOSLTVCKEVTGNYGVK	309
DB	301	IATPY--YLQSGTMSRGRNSPRAAPRDVAVPQIDG-----ETLLSMCQRVTNGYSHVK	352





QY 1051 EDEHLERTLEONKGMKMKKEK 1072  
Db 1014 LOSQLEORQOOSQOLSEKQO 1035

RESULT 3

528381  
utrophin - human  
N/Alternate names: dystrophin-related protein  
C/Species: Homo sapiens (man)  
C/Date: 17-Apr-1993 #sequence\_revision 03-Oct-1995 #text\_change 09-Jul-2004  
C/Accession: S28381; S28914; S03966  
R/Tinsley, J. M.  
submitted to the EMBL Data Library, November 1992  
A/Reference number: S28381  
A/Accession: S28381  
A/Molecule type: mRNA  
A/Residues: 1-3433 <TIN1>  
A/Cross-references: UNIPROT:P46939; UNIPARC:UPI0000137E6A; EMBL:X69086; NID:G34811; PIDN  
R/Tinsley, J. M.; Blake, D. J.; Roche, A.; Fairbrother, U.; Rise, J.; Bych, B. C.; Knight,  
Nature 360, 591-593, 1992  
A/Title: Primary structure of dystrophin-related protein.  
A/Reference number: S28914; MUID:9306045; PMID:1461283  
A/Accession: S28914  
A/Molecule type: mRNA  
A/Residues: 27-246;2839-3343 <TIN2>  
A/Cross-references: UNIPARC:UPI0000173E73; UNIPARC:UPI0000173E74; EMBL:X69086  
R/Love, D. R.; Hill, D. F.; Dickson, G.; Spurr, N. K.; Bych, B. C.; Marden, R. F.; Walsh, F.  
Nature 339, 55-58, 1989  
A/Title: An autosomal transcript in skeletal muscle with homology to dystrophin.  
A/Reference number: S03966; MUID:89238543; PMID:2541343  
A/Accession: S03966  
A/Molecule type: mRNA  
A/Residues: 2944-3433 <LOV>  
A/Cross-references: UNIPARC:UPI000000673; EMBL:X15488; NID:G30933; PIDN:CAA3515.1; PID  
C/Comment: This protein is found primarily at the neuromuscular junctions in adult muscle  
d regenerating muscle.  
C/Genetics:  
A/Gene: GDB:UTRN; DMDL  
A/Cross-references: GDB:119851; OMIM:128240  
A/Map position: 6q24-6q24  
C/Superfamily: dystrophin; alpha-actinin actin-binding domain homology; spectrin/dystrog  
C/Keywords: actin binding; cytoskeleton; leucine zipper; membrane-associated protein; mu  
F/30-248/Domain: alpha-actinin actin-binding domain homology <ACT>  
F/308-417/Domain: spectrin/dystrophin repeat homology <SP01>  
F/418-526/Domain: spectrin/dystrophin repeat homology <SP02>  
F/528-637/Domain: spectrin/dystrophin repeat homology <SP03>  
F/638-685/Region: hinge  
F/686-796/Domain: spectrin/dystrophin repeat homology <SP04>  
F/964-902/Domain: spectrin/dystrophin repeat homology <SP05>  
F/966-1013/Domain: spectrin/dystrophin repeat homology <SP06>  
F/1015-1122/Domain: spectrin/dystrophin repeat homology <SP07>  
F/1124-1230/Domain: spectrin/dystrophin repeat homology <SP08>  
F/1232-1334/Domain: spectrin/dystrophin repeat homology <SP09>  
F/1339-1450/Domain: spectrin/dystrophin repeat homology <SP10>  
F/1451-1544/Domain: spectrin/dystrophin repeat homology #status atypical <SP11>  
F/1543-1649/Domain: spectrin/dystrophin repeat homology <SP12>  
F/1651-1755/Domain: spectrin/dystrophin repeat homology <SP13>  
F/1856-1973/Domain: spectrin/dystrophin repeat homology <SP14>  
F/1975-2081/Domain: spectrin/dystrophin repeat homology <SP15>  
F/2083-2188/Domain: spectrin/dystrophin repeat homology <SP16>  
F/2227-2333/Domain: spectrin/dystrophin repeat homology <SP17>  
F/2335-2440/Domain: spectrin/dystrophin repeat homology <SP18>  
F/2442-2556/Domain: spectrin/dystrophin repeat homology <SP19>  
F/2558-2688/Domain: spectrin/dystrophin repeat homology <SP20>  
F/2690-2797/Domain: spectrin/dystrophin repeat homology <SP21>  
F/2798-2869/Region: hinge  
F/2812-2849/Domain: ww repeat homology <WW1>  
F/2837-3117/Region: cysteine-rich  
F/3263-3284/Region: leucine zipper motif  
F/3328-3349/Region: leucine zipper motif

Query Match 5.9%, Score 326.5; DB 1; Length 3433;

Best Local Similarity 19.8%; Pred. No. 6.4e-06; Matches 197; Conservative 159; Mismatches 336; Indels 303; Gaps 37;

QY 286 RKPNASGLTWCKEYVKNRGVYKTNFTTSMNRGSGFCALIHFRPDLIDKSLNPODI 345  
Db 148 QQTISEKILSWRQTPYRSQVAVNLFTTSTWDTGLFANVLRHKKEDFLSWDKVYKMS 207  
QY 346 KENKKAIDGFAS-IGISRLLEPSDWVLAIPDLITWYLYOI- 388  
Db 208 IERLEHAFSKAQYTLGIEKLLDPD-VAVLPPDKSITWLTSLFEVLPOQVITDAIREV 266  
QY 389 ---RAHFSQELNVQIENSSKSTYKVGNYETD---TNSVDQEFYALSDLKBBP 441  
Db 267 ETLPRKYKCECBEBALINQSTAPEEHESSPRAETPTVTVDMDLDSQYALBEV- 321  
QY 442 LQEPISGAVDFLGODSVFVNDGCVGSESEHQPDDHLSPTAS- 488  
Db 322 LTWLISADTFQEDD---ISD-VEEVKQFATHEAFMEELTAHQSSVGSVLQAGNOLI 377  
QY 489 CRATKSPTEPQKSG- 528  
Db 378 TQGTLSDEEFETQEQWTLNARWEALRVSMRQSLHD-VLMEIQKKQLQSLSMVLT 435  
QY 529 LGKKRLKAETLBSLSD- 556  
Db 436 LTERRIQMETCPLDVDVKSLOKLEBKSLQSDLEARQVNSLTHWVYVVDENSGESA 495  
QY 557 PFL- 599  
Db 496 TALIEDQLOKIGRWTAVCWTEBRMVRLOEINI-LWELBLE- -CCLKAMVLEK 548  
QY 600 KTSLSPTSKLGYSRDLDAKKKHAISLRQTESPDADRTLLNADSSKIVQ- 652  
Db 549 BEALNKQTSNFKDQKSLSVRLALKE- -DMEKROTLL- -DLSIGQDVGLD 602  
QY 653 ---HRLSROBELKER- 678  
Db 603 NSKASKKINDSEELTRQWDSLVQRLEDSNQVTAQAVKGMQIPQKDLLETVRVAEQA 662  
QY 679 LKAGNKNTWTAIPFCNRQLSDQDERRRQLERRAQL- -IAEASGYMAGSLPYGE 735  
Db 663 ITTKSKOELPPPPPKRQI- -HYDIEAKKFPALISLELNLWKMTALQTEIEIKYMK 720  
QY 736 MA-AEKIKERSKASGENDNIBIDPNEIEPGFVVGSGDELTLNENDLDPQNSKLVLD 794  
Db 721 MQDTSEKKKKA- -LEKQERLIR- -ADBLNQTGQLVEMQKSGILPTE 767  
QY 795 KLKLELVQPVANSPSSAAQKAVTESSEQDMKSGTEDLTERLQKTERFRNPFVPSKD 854  
Db 768 EIKNVLE- -KVSSEMKNVSGHLEDLERK- -IQAQED 799  
QY 855 STYKQYQLOFSFOYIENRPMKORSIQEDTK- - 892  
Db 800 INNVFKQDLLEKVIKTKKEWVGHGTSISRSQSLPALKSCQRELTNLGLHPKIMAR 859  
QY 893 AATTEQRKSEBDEVANKF- -KDSQYVGEALALENQKQIDIRAA 938  
Db 860 ASGALMSQPSADPFVGRGDFLGRYQAVQAEVDEQEH- -LENELKQOPGHAY 912  
QY 939 LVEKRLLEYLMDTGRNT- -EEASMMQEWFMVVKNAALIRRM 978  
Db 913 L--ETLKLTDVANDSENKQVSLANTLNDLAKYKALQEKKTIDE- -ILNQPALHKL 968  
QY 979 NQSLLEK- -EHDLERRYELLNBLRAMLAIEDWQKTBQAKRRBQLLDLDELV- 1028  
Db 969 EETKALEKVNHPDVEKLYKQEPDVG- -KWKRLKVLVSKDILLBEIALTLRAFEA 1023  
QY 1029 --ALVNRKDALVDDLAQEKQAEEDBEHLERTLEQ 1061  
Db 1024 DSTVIEKMDGVXDPLMKQQAQGDAGLRQLDQ 1058

RESULT 4

T34418  
 hypothetical protein F12F3.3 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T34418  
 R:Fulton, B.; Wohlmann, P.  
 submitted to the EMBL Data Library, July 1998  
 A:Description: The sequence of C. elegans cosmid F12F3.  
 A:Reference number: Z21521  
 A:Accession: T34418  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-3488 <FUL>  
 A:Cross-references: UNIPARC:UPI000017B8E6; EMBL:U80022; P1DN:AC25885.1; GSPDB:GN00023;  
 A:Experimental source: strain Bristol N2, clone F12F3  
 C:Genetics:  
 A:Gene: CESP.F12F3.3  
 A:Map position: 5  
 A:Insertions: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match 5.9%; Score 324.5; DB 2; Length 3488;  
 Best Local Similarity 19.1%; Pred. No. 8e-06;  
 Matches 245; Conservative 215; Mismatches 429; Indels 397; Gaps 54;

25 VVSAAQFSLSCITFLREGKATDEDMOSLASVSMKQADIGNLDP---FEEDNED-DEENR 80  
 315 VAGEAWCFSDVVVHMGSE-BSRDD-----KSVDEVNDSTVLEKKDGDGDKSK 361  
 81 VNQEEKAATKTEILINKLFLDEAKKOLATVNSPPDDPAEELNPGDPDSEPTETAS 140  
 362 PKYKKKTIKKKE--TPSEQVTALEPEQOKISEVDVQSVATEVEGAKKKPAEKP--TDLSK 419  
 141 PRKTEDSFVNSVNPPEKAVQTPOLNPPDEEAFVITKD-----SPPOSTKKNR 191  
 420 AKKDSK-----KKSDEPEASTEEKSTTEKPTNDKTSKSAEKKTVK 461  
 192 PVDMSKLYADSSKTEEBELDSNPFYKSTPPNNLVNPPVOELETERVKKKAP--- 247  
 462 PKKEVVGKPLEAKCPEDKKDASQPSKSSKSSPP-----TDGKKKKQIKALF 509  
 248 APPVLSPTKGLVLENTVSAKGLSTSPKSPSPVYGRKPN--SQSLVWCKEVTKNR 306  
 510 IPDELSRRFG-----DPTVMSEFTNITTTIRGEGSADAKTPLY--EPLSASV 555  
 307 GVKITFTTSMRNGSFCALIHFRPD-----LIDVKSINPQD-----IKE 347  
 556 SMKVFPTLVESAKEKAFSPKRSRSETDDSKRKEGPRPAKSEKDEVTAEKOSTALE 615  
 348 NNKKAVDGFASIGISRLR--PSD--MYLLAIPDKLVTWTVLYQIRAHFSGQEL--NVV 400  
 616 SKKKEVDE-----SKISFQPSDKNKSIVGVPEKA-----AGPETKKDVS 656  
 401 QIEENSSKSTYVGNVETDTNSSVDQEKFYAEISDLKREBELQOPISGAVDFLSQDSVF 460  
 657 ERIEVPKKTITIKKTEKSD--SSISQKSNVLKPAD-----DDKSKSDV- 698  
 461 VINDSGVSESHQRPDDHLSPTSPYCRKTSQDTEPKSQSSGRTSGSDPGICNMT 520  
 699 -----TDSKTKTETDQTKVATDS-----KLEPAADTTKQILETVVDDK----- 737  
 521 DSTOAGVLLGKRLKAEITLSDLVSDKKDMSPPFI-----CEETDEQKQLTL----- 571  
 738 -----SKKVLKKKT-EKSDSFISQSR--TPPVVEPTKPAESEKQAEVNVAKK 785  
 572 --DIGSNLEKELIENSRL--ECSRDPSPPIKTSLSPTSCLKGYSDLDLAKKQAS 626  
 786 QKEVNDNLKREAEVAAKIADEKLKIEAANIKKA-----EVEAAKKQKSK 832  
 627 LAQTESDPP-----ADRTTLNHADHSKTYOHLISNOEELKEKARVILEQARDAA-- 678  
 833 DEQLKLETVVSKSAERKLEKQAIKKAELADVVKQKEINENK--LEAAKKSAD 890  
 679 -----LKAQNKHNTNTATPFCNRQL-----SDQDQE 704

Db 891 KKLSESAKSKKVSSEYKPEGEKKTAKGETVQVSEFPSTKITIDTKDVGATEPADE 950  
 Qy 705 ERFQL-----REPAPQLAE--ARSGVQSE-----LPSYGEAAAEKTKERSQASGD 750  
 Db 951 TPCKKTIKKTKESDSSISQKATDESEKSKQKODEPTKPAVSEFQMTVEADSKSKQKE 1010  
 Qy 751 ENDNITID-----TNEBIPGVVGGSGELTLENDLTPBEQNSGLVD--LKI----- 796  
 Db 1011 TDEKLDLDEIAKTAQKQADEKSKLDQEKIKKVSSE--DAARKEKELNDLKLESEIATK 1069  
 Qy 797 -----KCLEVOPQVANSPPSAQAQVTSSEQ-----DKMSGTEDLRTE----- 836  
 Db 1070 KASADLKLEBQQAQKAAVEAAKQKQKEDQLKDTEAASKKAAAEKLEKQAIKK 1129  
 Qy 837 -----RLQKTERFRANVPSKDSVTAKTQLOQSQYIENRPEMKRORSIOEPTKGN 889  
 Db 1130 AAGADAVKQKQKELDE--KNLLEANKKSAAKTLK-----IEESAAKSKQVVEBQAKLDA 1181  
 Qy 890 EEKKAITETQKPSDE-----VLRKGFK-----DTSQYV 920  
 Db 1182 QTKAKTAKEQTKLEKDEKSTKSESSEKTYDEKPKKVLKKTKEKSDSSISQKSEISKTV 1241  
 Qy 921 GELALENE-----OKQIDTPRAL-----VEKRLRYL--MDTG 951  
 Db 1242 ESAGSESESTQVADAARQKQKTEDEKQLBAKITAKKSADESKLEAEKTKKAAVEAA 1301  
 Qy 952 RUTTEERAMQEWFMVLRKAL-----IRRMQSLLEKHDLEFRRYELN--- 998  
 Db 1302 KQKKEDEQLKLDTEAASKKAAAEKLEKQSHIKKAAEVNDVVKQKELEEKORLESEA 1361  
 Qy 999 -----RELRLMALIEDQKTEAQKRREROLLDELVALVNRDLVLRDLQ 1044  
 Db 1362 TKKADAEKLEBQKKAALAILALIEIQKQEKQALQSQSLDEBA---KQSAEKQKLESE 1417  
 Db 1418 TSKQTEAPKESVDEKPKKQVYLRKK 1443

RESULT 5  
 PADOAA  
 alpha-actinin - slime mold (Dictyostelium discoideum)  
 C:Species: Dictyostelium discoideum  
 C:Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 09-Jul-2004  
 C:Accession: S00103; A29006  
 R:Noegel, A.; Witke, W.; Schleicher, M.  
 FEBS Lett. 221, 391-396, 1987  
 A:Title: Calcium-sensitive non-muscle alpha-actinin contains EF-hand structures and high  
 A:Reference number: S00103; MUID:87304850; PMID:3622778  
 A:Accession: S00103  
 A:Molecule type: mRNA  
 A:Residues: 1-862 <NOE>  
 A:Cross-references: UNIPROT:P05095; UNIPARC:UPI0000125099; EMBL:Y00689; NID:G7177; PIDN:C  
 J. Cell Biol. 103, 969-975, 1986  
 A:Title: Studies on the transcription, translation, and structure of alpha-actinin in Dic  
 A:Reference number: A29006; MUID:86504574; PMID:3745276  
 A:Accession: A29006  
 A:Molecule type: DNA  
 A:Residues: 92-359, 'P', 361-500, 'T', 502-505 <WIT>  
 A:Cross-references: UNIPARC:UPI000016898B; EMBL:X04324; NID:G7202; PIDN:CAA27855.1; PID:C  
 C:Suprafamily: alpha-actinin; alpha-actinin actin-binding domain homology; calmodulin rel  
 C:Keywords: actin binding; calcium binding; duplication; EF hand; homodimer  
 F:21-236/Domain: alpha-actinin actin-binding domain homology <SP1>  
 F:266-377/Domain: spectrin/dystrophin repeat homology <SP1>  
 F:386-493/Domain: spectrin/dystrophin repeat homology <SP2>  
 F:502-607/Domain: spectrin/dystrophin repeat homology <SP3>  
 F:616-717/Domain: spectrin/dystrophin repeat homology <SP4>  
 F:730-762/Domain: calmodulin repeat homology <SP1>  
 F:766-798/Domain: calmodulin repeat homology <SP2>

Query Match 5.8%; Score 321; DB 1; Length 862;

Best Local Similarity 23.3%; Pred. No. 1.9e-06;  
Matches 187; Conservative 98; Mismatches 262; Indels 256; Gaps 35;

```
QY NASOSLWCKEVNTKNGVKTNTFTSMRNGLSFCALIHHPRLIDVXSLNPDIEN 348
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 289 SAKELLLMCRKTEGYVKNVGNHTSPQDLACALIHKRPDLINPDSLMDKXGN 197
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 349 NKKAVD-GFASIGISRLLEPSDMVLLAIPDKLTVMYTYLQIRAHPSGOELNVQIEKSS 407
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 198 LQALFDIAEKELDIPKMLDVEDMDLVAPDERSVWITYAQYTHHSAS-----RKAETAG 252
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 408 KSTYKGVNGETDYNSSVDOEKRYAELSD-LKREPELQPIGAVDFLSGDDSVFVNDSCV 466
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 253 KQVGVN-----LDTFMLLEQTK-----SDYLKRAVELVQMIN-----DKQASLESDF 295
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 467 GES-ESHEQTPDDHLSPTASPYCRKTSPTPEQSGSSGRTSGSDPGICSNITDIOA 525
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 296 GDSIESVGSFNNAH-----KEYKTEKPPKQGEVS-----ELEA 329
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 526 --QVLLGKGRLLKKAETLESLDLYVSDKKGMSPPETC-----BERTBOKLQTLIDISNL 577
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 330 IYNELQTKLILKE-----PFAVAPAGLTPNEID-----STWSAL 364
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 578 EKEKLENSRSLSCRSDESPPI-----KKTSLSPSTKLGYSYRDV-DLAKKXHA 625
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 365 EKAQEHAEALRIELKQKCTAVLLQKYNRLIKLENNATTKSVYLGSENETGDSITAYQA 424
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 626 SLRGTE-----SDPD-----ADRTTLNH-----ADHS 648
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 425 KLKMLAEVDEGCQSLSEGSNDLLSLIQLTELNVGVPELTERKOTPEAQQWGVKSSA 484
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 649 KIYVHRLISROBELKERARVILLEQARRDAL-----KAGNKH-----NNTATPFCNRQ-- 697
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 485 ETYKNTLAELEBRLOKLEDSLVEFPAKRAQLNWMLEADHDVFPDFINVSQVGOEIOEK 544
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 698 ----LSPDQD-----BERRQ 709
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 545 FDAFLHDQSQFALEALALATQQLRELGRSENDYSVSYDELAKMNNLLAGIEBRVQ 604
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 710 LREBARQLIABARS-----GVKMSLPSYGEAIAEKLKERSKASGENDN---LEID 758
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 605 L--ANELLTYQNNVLCQSPSVKANEISDYRVYTLDAISQNTSDPOEQALNTRAITTA 661
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 759 TNEBIPGFFVVGGBELTNLENDL-----DTPGNSGLVDLKLK-----KLEVQPOVA 807
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 662 HAERKPE-----LDELITTSQLEBAQVNDKHTQHSLSLKLKMDKLTAKKNEGV 715
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 808 NSPSSAAQ-KAVTESSEQDMKS-----GTEDLRTERLQ-----KTERFRNPV 850
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 716 BGEILAQULGVTALELSEFACSHPKDNDKLNRLREFSSCLKSIGDELTEQLNQYI 775
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 851 PSKQSTVRKTQLOFSQYIENRPEMKQRSIQEDTKG---NEKAAITETQRPSEDE 906
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 776 SKITPDNGT--ISFEERIDYVSSRKGTDSVESTKAFAKVAEDKDPITBAQ----- 826
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 907 VLNKGFQDTGQ--YVGEIATLE 927
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 827 -IRRAISDSKQIDYLLASMPAVE 848
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
beta-spectrin - mouse
C/Spect: Mus sp. (mouse)
C/Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 13-Aug-1999
C/Accession: I52577
R/Bloom, M.L.; Birkenmeier, C.S.; Barker, J.E.
A>Title: Complete nucleotide sequence of the murine erythroid beta-spectrin cDNA and the
A/Reference number: I52577; MUID:9403578; PMID:8219233
A/Accession: I52577
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
```

A/Residues: 1-2128 <RES>  
A/Cross-references: UNIPARC:UP10000029884; GB:S66263; NID:9440899; PIDN:AA828600.1; PID:  
C/Genetics:  
A/Gene: Spnb-1  
A/Introns: 418/3; 1742/2  
C/Superfamily: spectrin beta chain; alpha-actinin actin-binding domain homology; spectri  
C/Keywords: actin binding  
F/53-271/Domain: alpha-actinin actin-binding domain homology <ACT>  
F/1787-1892/Domain: spectrin/dystrophin repeat homology <8P>

Query Match 5.8%; Score 319; DB 2; Length 2128;  
Best Local Similarity 19.7%; Pred. No. 7.3e-06;  
Matches 221; Conservative 160; Mismatches 382; Indels 358; Gaps 44;

```
QY 285 GRKE-NASOSLWCKEVNTKNGVKTNTFTSMRNGLSFCALIHHPRLIDVXSLNPQ 343
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 169 GREQSAKDALLMCCQKTAGYFHVANNYTNFTSSWKDGLAFALIHKRPDLIDPDKLDS 228
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 344 DIKENKKAVD-GFASIGISRLLEPSDMVLLAIPDKLTVMYTYLQIRAHPSGOELNVQ- 401
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 229 NARHNLHAPVARQGLIPLDPEQ-VFTEPDEKSIITVVAFFHYFBSKMYLAVG 287
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 402 -----IEE-----NSSK-----STY 411
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 288 KRVGKVIDHALETEKMIKYSGLASDLLTWIQTIVLSNRKFRANSLGVQQLQAFSTY 347
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 412 KV-----GVYF-----DTNSSVDOEKRYAELSD-LDLKR-----PELQOP 445
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 348 RTVEKPPKFOEKGMLEVLLFTIQSRMRANQKVTPPHDGKLVSDINRAMESLEAEVQRE 407
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 446 ISGAVDFLSODSVFVNDSCGVG--SESEHQTPEDDHLSPTASPYCRK-----SDTE 497
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 408 LALSELIRQE--FDKKAARETWLNQGLVQDNPFGYDLAVEAKKHEALETITA 464
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 498 POKSQ-----QSGRTSGSDPGICSNYDS-----TQAOVLLGKGRLLKAELESLD 544
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 465 AYERVRALIEDLAQLEKENYHDKRIIARDNDILRMSYQLBELRSRQLEBATLALQ 524
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 545 LYVS-----DKKMSPPICEETDEQKLTQTLIDGSLMEKLENSSLSEKSPSESTY 599
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 525 LFQMLHSIDMDEIKHILISAEFGKILLEVEDI--LQKHKLEBA-DIALQGDVKAIT 580
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 600 KTSLSPTSKLGS-----YSRDLDAKKKHAELRQTES----- 632
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 581 AATIQFABGKYQPCQDVYQDRVSHLEQCFSELSNPAAGKQLEBSKLMKFEWED 640
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 633 -----DPADRTTLNHADHSSKIVQHRLISROBELK----- 664
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 641 AEWIKERKEQIYSLDYGKDLTSVLLQKHKHAFEDELRGIDAHLKQIPQADDMVAQK 700
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 665 -----PARVILBEQARD--AALKQGNKHTNTATPF-----CNROLSD 700
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 701 FGHQPIETRVKVASAQMDHKEILAFAKKQLQDAENFFQFGADDDKAWLQDAHRLISG 760
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 701 Q---ODEBRRQLAERARQLIABARSGVKNSELPSYGEAIAEKLKERSKASGENDNIEI 757
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 761 EDVQDDEGATRALGKKHKEFLBELBSRGVWE--HLEHQAGPPEERFRSPDVTNRLOA 817
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 758 -----DTNEBIPGFFVVGGBD-----ELTNLENDLDT 784
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 818 LRKLYQGVLLQALRGHKLQLEALDLYVPGESDCELMWTEKGMLOMDIPNLEDELEV 877
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 785 PEONSKLVLDLKLKLELVQPOVANSPESSAAQKAVTESSEBDMKSGTEDLRTERLQKTER 844
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 878 VQHRFDILDOBMKTLMAQIDGV-----NLANNLVESGHP--NSGEVQYQDRANLKXWQA 930
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 845 FRNPVYVSK---DSTVKK---TQLQSFQYIENRPEMKQRSIQEDTKKGNNEKAAITE 897
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 931 FOAVVSGORAEVABALVNNYTCVDCSTSKMIDKTK-----VVESTKQLQGLQAGVIA 984
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 898 TORKPS--EDEVLN-----KGFQDTSQYVVGELIATLENBOKQ-IDTPAALVEKRLRYLMD 949
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 985 IQRKLSLEBRDYLAIRVSALEBRSGY-----LMESHPEQKBDIGQQAQVYKMLKQLQD 1040
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

QY 950 TGRNTE-----BEEANMOE-----YFNLVNKNA-----LIRRNQDLSLEK 986  
Db 1041 ALQGOEUSLGEASKDQAFLDLDDFKALNSAQKVASSEDMPEBSLPEAREQLLOQHAATKE 1100  
QY 987 EHDLER-----RYELLNRELRAM-----L 1005  
Db 1101 EIDAHRDDYHRVKASGEKIVIEGQTDPPDYQLLGGRLGEGLDTDMDALRWMSRGNTLITQCL 1160  
QY 1006 ALEIMQKTEAC----KREQLL-----DELVALVNKRDALYRDL 1041  
Db 1161 GFQGFQKQAKAEAILNSQOEYTLAHLEPPDSLAAAEAGIRKFEDFLVSMENNRDILSPV 1220  
QY 1042 DAQEKQAEEDHELRTE-----QNKGMAKKEKCVL 1075  
Db 1221 DSGNLTVAEGNLVYSNKTIEMKEVQLIEDRRHKKNKEQAQETVL 1261

RESULT 7

667593

transport protein USO1 - yeast (*Saccharomyces cerevisiae*)

N/Alternate names: protein D2552; protein YD1055W

N/Species: *Saccharomyces cerevisiae*

C/Date: 12-Jul-1996 #sequence revision 12-Jul-1996 #text\_change 09-Jul-2004

C/Accession: S67593; A38455; S30782

R/Bioecker, H.; Brandt, P.

Submitted to the Protein Sequence Database, July 1996

A/Reference number: S67587

A/Accession: S67593

A/Molecule type: DNA

A/Residues: 1-1790 <BIO>

A/Cross-references: UNIPROT:Q07380; UNIPARC:UPI0000052PC0; EMBL:E74106; NID:g1431058; PMID:2010462

A/Experimental source: strain S288C

R/Nakajima, H.; Hirata, A.; Ogawa, Y.; Yonehara, T.; Yoda, K.; Yamaeaki, M.

J. Cell Biol. 113, 245-260, 1991

A/Title: A cytoskeleton-related gene, USO1, is required for intracellular protein transport

A/Reference number: A38455; MUID:91185402; PMID:2010462

A/Accession: A38455

A/Molecule type: DNA

A/Residues: 1-389, 'TA', 392-724, 'S', 726-1790 <NAK>

A/Cross-references: UNIPARC:UPI0000137B56; GB:X54378; NID:g4777; PIRN:CAA38253.1; PID:g4777

A/Note: the authors translated the codon ACT for residue 768 as Ile

R/Hoebster, M.K.; Herman, D.J.; Bendall, C.M.; McClellan, M.; Tao, N.; Kendrick, K.E.

Submitted to the EMBL Data Library, February 1993

A/Description: An integrin analogue in *Saccharomyces cerevisiae*.

A/Reference number: S30782

A/Accession: S30782

A/Molecule type: DNA

A/Residues: 71-846, 'E', 848-923, 'K', 925-1552, 'I', 1254-1318, 'V', 1320-1460, 'S', 1462-1580, 'S'

A/Cross-references: UNIPARC:UPI000017B38F; EMBL:L03188

C/Genetics:

A/Gene: SGD:USO1, INT1

A/Cross-references: SGD:S0002216; MIPS:YD1055W

A/Map position: 4L

C/Keywords: coiled coil; transmembrane protein

F/336-342/Domain: transmembrane #status predicted <TM1>

F/394-410/Domain: transmembrane #status predicted <TM2>

F/617-633/Domain: transmembrane #status predicted <TM3>

```

Query March      5.8%; Score 318.5; DB 2; Length 1790;
Best Local Similarity 19.1%; Pred. No. 6.2e-06;
Matches 217; Conservative 220; Mismatches 392; Indels 307; Gaps 47;

QY      43 KATDEDMQSLASLVMKQADIGNLDPEEDNEDDENVKQAEKATITELINKLNFPLDE 102
      | : : : : : : : : : : : : : : : : : : : : : : : :
Db      726 KISFEVEKIKQROCTKXKGRITSL--QTRETSHEINLITE-----KLIALTVEHKELDE 776

QY      103 AEKOLAVVNSPFDDPDPAELNFGDPDSEPIETETASPKRTEDSFYNSGVNPEKQVPT 162
      | : : : : : : : : : : : : : : : : : : : : : : : :
Db      777 KYQILNSHSLSKENKFSILLETLEKNVDSLDDEMQLDVLVETKO-----KEMQTA 826

QY      163 --QVLNPFDEBEAVTILKDSPPQ--STKRKNIRPVD--MSKYLVA-----DSKTE 207
      | : : : : : : : : : : : : : : : : : : : : : : : :

```

Dh	827	LLEKYSTIHKQDSDIKTLEKGETLISQCKKAEIDGINKKQDOLFALSREMOVAEBNCKNL	886
Qy	208	EEELDESNPTEPPKSTPPPPNNLVNVOELETERRVRKKAAPAPVLSPPKGVLENENTVSAG	267
Dh	887	QEKOKSNVNNHOKETKSLKEDIAAKTEIK-----AIENL-----	922
Qy	268	KOLSTSPKSPPIPSVPLGRKPNASOSLWCKEVTKNRGVKTINFTTWSRNLSPFCAIL	327
Dh	923	-----EBMKIQNNLSAE-----K	936
Qy	328	HHFBDLIDYKSLNPODIKENNKAYDGPASIGISHLSPSDWVLLAIPDKLTWMTVLYQ	387
Dh	937	BHISKELVEYKS-----RFQSHDVLVAKLTETKLSLANNYK	972
Qy	388	-IRAHFSGOELNVNOIEENSCKSTYKGVNEYETDMS-SYDQEKFYAELSDKREPE-LOQ	444
Dh	973	DMQAR---NESLIKVERKSNKSSIQLSNLOKNIKMSQSEKFNFOIERSIKNIQOLKK	1020
Qy	445	PISAIVDFLSODSYFVNDSDGVGSESEHQETDDHLSPGTASPYCARTSDTEPORSQOS	504
Dh	1030	TISDLEQ---TKREITISKDSDSODEYBSQLSLKEKLETFY-----TANDEVANKISE-	1077
Qy	505	SGRITSGSDDPGICSNSTDSOQAVLLGKKRLLAET-LLESLDLYVSDKKKMDSPPICBET	563
Dh	1080	-----LTKTREELAEIAYANULKNLELFTJLFTSEKALKEVENE-----EHL	1122
Qy	564	DEQICQTLDDISGNEKEKYLENRSJ-ECHSDPESPPIKXTSLSPSTGLGVSYSRDLLA--	620
Dh	1123	KEBKIQ-----LKEKATETKQOANSLCANLESLEKHE-----DLAQ	1166
Qy	621	-KQGHASIRQTESDDADRTTNHAHDSKIQYHRLLSHQEELKEBAYVLLQOARDAAAL	679
Dh	1161	LKTBEOQANERQYNEISQUNDEITTSIQQNEBESIKKGNDELBGVYKAMKTSBEOQNL	1222
Qy	680	KAGHKHNTNVTAPFCNRQLSD--QODEBRRLREBAROLIAEARSQVMSBLP---SYG	734
Dh	1221	K-----KSIDALNLOIKELKKNQETBASLBSIKVSESET---VKIKELQDCNFK	1277
Qy	735	EMAEKLEKERSKASDENDN-IEID-TNBEIPEGFVGGG-----DELTNL-----E	779
Dh	1271	EKEVSELBDKIKASEDKNSKYLELOKESBKIEBELDAKTEBELKIOLEKITTNSLKAKEKE	1333
Qy	780	NDLD---TPRONSLVDLKLKGL-----LEVQOVANSPSSAOKAVTES-----	821
Dh	1331	SELKRLKTTSEBRKNAEQLKLANEIQIKQAEPKEKLANEGSSTITQETSEKINTL	1397
Qy	822	-----SEODMKSGETDLTERLOKTERFRANFVPSKDSVTK	859
Dh	1391	EDELIRLQNEBELKAKEIDNTSESEKVSLSNDELLEBKONTIKISLODEILSYKDKITRN	1457
Qy	860	TQ-IQSPQOYIENRPMKROBSIQODTKKGNKAKAIYITQKRPSPSDVYLVANKPQTSQY	918
Dh	1451	DEKLSTIR--DNKKDLB--SLKQOLPAAQSSKAKVEGGLKLLEBSSSEKAELEKSKR	1507
Qy	919	VVGEI-AALENBOKOIDTRAAVLVEKRLRYLMDQGTGNTBEEEAAMQEWFMVLVKNKALLIR	977
Dh	1506	MMKLTSEITSENETLKSMEYIRKSDKELBSSKKSABEIDKNLQ-----HEKSLIR	1557
Qy	978	NNQUSLEKEHDLERRYELNRELPAMLAIE-----DMQTEAQKREBOLL--DELIV	1022
Dh	1560	IN-----ESEKDIR-----ELKSKIRIEAKSGSELLEVQELINNAQEKIRINARENT	1607
Qy	1029	ALVNRKDALVYDLDQEKQAE-----EEDD-----HLERTLQONKGMKAKKEK	1072
Dh	1607	VKSKLEBIEBEL--KDKQAEIKSNOEKELITTSRLKELBQSLDSTQQAQSSBEER	1660

```

RESULT 8
A46147    spectrin beta chain - fruit fly (Drosophila melanogaster)
C1:Species: Drosophila melanogaster
C1:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C1:Accession: A46147; A33657

```





```

Db      401 ENRQVAIIAELROKLEQLAQRFHKK-----AKLRDSWLRASQVVLSEMEHGRSASQV 455
Qy      531 KKRLLKATLELSLVDKDKKMSPPFICEETDEQKLTIDLSNLEKELNRSRLLEC 590
Db      456 EKLTKQQAISTDILABEDRFKMLTA--MCNELCTKHESDVKRGEMERETIDMTQLLT 513
Qy      591 RSDPESPICKTSLSPSTKLGYSYRDLAKKKHSAIRQTESDDPDAARTTLNHAHSSKI 650
Db      514 LLEQ----RRBALMSLNDL--MSLRDIDTLNSNELYSLEPAVRNRDVCKHLIGVEDLLGK- 567
Qy      651 VOHRLLSRQ-----BELKERARVLL-----QARRDALAKGKNTNTATPF 693
Db      568 --HDLLDAQINAGSLSLKLSQSANNYIRHKEQPDVLRKLDDEVTA-----QYNTLVEL 620
Qy      694 CNRQLSDQDQEBERRQLRERARQLIARSGVMSLEPSYGMA-----AEKLERSKASG 749
Db      621 C-----RSRRGLERARSLFQVQDHER-----EMWMLAEKELCTTALNSG 662
Qy      750 D-----ENDNIEIDTNEBIPGFVVG-----GDELTNLENDL----- 782
Db      663 DISAVPQTLLYKXVEMQTHMARSKGMIAGGERLVQNGSKEDIQRLLTQNMHRWERL 722
Qy      783 -----DTPRQNS-----KLVDL----- 794
Db      723 RVAVDALGNWLSERARHAQOYFQDANBARASWIREKMPLVKSDDLGRDEGAASLLQRHARL 782
Qy      795 -----KLKLLLEVOPOVANSPPSAQAQVATESSEQDKSTEDLRTERLOKTERF 845
Db      783 BEETIRAYKSDISIRLEWQSOLAN--SAFTTATTSQVQ-----TEBVNPQVEMSYNTE 835
Qy      846 RNPVFSKDSVTRKTOLOSFSQYIENRPE-----MKRQR-----SIQ 883
Db      836 GNGMRYSKGBVLLALF-----KSTPEWMLAKKDGTEGYVPANYCKIVPGETVTVQ 887
Qy      884 DTKK-----GNEEKA--ITETQRKPSDEVLNKGFDTSQYVVG-----ELAAL 926
Db      888 TTQKTTTLEGNETKSSVADRHQKISNDYRELKRLADVRRLSDNIKILRFYRCDER 947
Qy      927 ENRQKQIDTBAALVEKRLRYLMDTGRYTEBEEBAMQEMFVLVKNKALIRBNQOLS--LL 964
Db      948 ERMAKEIEVSLA--DESPERHVAARFRKPKLEADMK-----TNGTQLKHINDIANDLI 1000
Qy      985 EKEHDLERRRYELNREIRAMLAIEDWQKTEAQKRREQLLDELVALVNRDALVRDLDAQ 1044
Db      1001 SEBHGSQRKIEVQHKNINM-----WNLERLRKQGRVLE-----ATER---VADPDTT 1047
Qy      1045 EKOAE-----EDHLEKTELEONKGMKKKE 1071
Db      1048 CESAREWMLSKFEQLDRNPVDVKSQVLERDLKPLEDKIALERK 1091

RESULT 10
T23630
hypothetical protein R31.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T23630; T24242
R:Kernaw, J.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19774
A:Accession: T23630
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-4101 <W12>
A:Cross-references: UNIPROT:O02425; UNIPARC:UPI000017CF3B; EMBL:Z81570; PIDN:CAB04608.1;
A:Experimental source: clone K12G11
R:Lennard, N.
submitted to the EMBL Data Library, July 1996
A:Reference number: Z19862
A:Accession: T24242
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

```

```

A:Residues: 1-4101 <W12>
A:Cross-references: UNIPARC:UPI000017CF3B; EMBL:Z75956; PIDN:CAB00130.1; GSPDB:GN00023; (
A:Experimental source: clone R31
C:Genetics:
A:Gene: CBSP.R31.1
A:Map position: 5
A:Insertions: 29/3; 89/3; 184/3; 312/3; 812/2; 1461/3; 1531/3; 2050/2; 2413/2; 2466/2; 3008/
Query Match 5.8%; Score 316.5; DB 2; Length 4101;
Best local similarity 20.3%; Match No. 2,1e-05;
Matches 204; Conservative 150; Mismatch 353; Indels 297; Gaps 41;

Qy      285 GRKNASOSLWVCKEYTKNYRGVKTINFTTSWRNGLSFSCALIHRRPDLIDYKSLNPOD 344
Db      190 GERKHAQDMLLWCGKRTAGYPNVRIENFTTSVRNGIAFVALHSHRPDLVDPRNLNPN 249
Qy      345 IKNNKKAYD--GRASIGISHLPEPSMVTALIDKLTMTYLYQIRAHSGQF----- 396
Db      250 HVDNLNHAEPVAKKEILAILDAED--VDVTRDEKSIITYVSLYYHHPAKQKTEMTGAR 308
Qy      397 --LNVV-----QIENSSKSTYKGVNYETDNNSSVDOEKFAELSDLKR----- 439
Db      309 RIANIVGKLWSTMEDDYEHIASLELNLIRVITKLESRRFPNSLNGKREBAKFNQFR 368
Qy      440 -----PELOQPSGAVDFLSQDDSVFVNDSGVSESEHQTPDDHLSPSTASPYCRTKS 494
Db      369 TSEKPKYKX--KGHEAL-----FTIQTAKMRSKQYQPPQGLFMHDIESAMAQLDVA 422
Qy      495 DTEPQ-----KSQSSGRTSGSDPGICSNSTD--TOAQVLLG----- 530
Db      423 ENRQVAIIAELROKLEQLAQRFHKK-----AKLRDSWLRASQVVLSEMEHGRSASQV 477
Qy      531 KKRLLKATLELSLVDKDKKMSPPFICEETDEQKLTIDLSNLEKELNRSRLLEC 590
Db      478 EKLTKQQAISTDILABEDRFKMLTA--MCNELCTKHESDVKRGEMERETIDMTQLLT 535
Qy      591 RSDPESPICKTSLSPSTKLGYSYRDLAKKKHSAIRQTESDDPDAARTTLNHAHSSKI 650
Db      514 LLEQ----RRBALMSLNDL--MSLRDIDTLNSNELYSLEPAVRNRDVCKHLIGVEDLLGK- 589
Qy      651 VOHRLLSRQ-----BELKERARVLL-----QARRDALAKGKNTNTATPF 693
Db      590 --HDLLDAQINAGSLSLKLSQSANNYIRHKEQPDVLRKLDDEVTA-----QYNTLVEL 642
Qy      694 CNRQLSDQDQEBERRQLRERARQLIARSGVMSLEPSYGMA-----AEKLERSKASG 749
Db      643 C-----RSRRGLERARSLFQVQDHER-----EMWMLAEKELCTTALNSG 684
Qy      750 D-----ENDNIEIDTNEBIPGFVVG-----GDELTNLENDL----- 782
Db      685 DISAVPQTLLYKXVEMQTHMARSKGMIAGGERLVQNGSKEDIQRLLTQNMHRWERL 744
Qy      783 -----DTPRQNS-----KLVDL----- 794
Db      745 RVAVDALGNWLSERARHAQOYFQDANBARASWIREKMPLVKSDDLGRDEGAASLLQRHARL 804
Qy      795 -----KLKLLLEVOPOVANSPPSAQAQVATESSEQDKSTEDLRTERLOKTERF 845
Db      805 BEETIRAYKSDISIRLEWQSOLAN--SAFTTATTSQVQ-----TEBVNPQVEMSYNTE 857
Qy      846 RNPVFSKDSVTRKTOLOSFSQYIENRPE-----MKRQR-----SIQ 883
Db      858 GNGMRYSKGBVLLALF-----KSTPEWMLAKKDGTEGYVPANYCKIVPGETVTVQ 909
Qy      884 DTKK-----GNEEKA--ITETQRKPSDEVLNKGFDTSQYVVG-----ELAAL 926
Db      910 TTQKTTTLEGNETKSSVADRHQKISNDYRELKRLADVRRLSDNIKILRFYRCDER 969
Qy      927 ENRQKQIDTBAALVEKRLRYLMDTGRYTEBEEBAMQEMFVLVKNKALIRBNQOLS--LL 964
Db      970 ERMAKEIEVSLA--DESPERHVAARFRKPKLEADMK-----TNGTQLKHINDIANDLI 1022
Qy      985 EKEHDLERRRYELNREIRAMLAIEDWQKTEAQKRREQLLDELVALVNRDALVRDLDAQ 1044

```



Db 1023 SEHGQSKIEVRQKIKIAM-----MDLSEKLRKRGGRLE-----ATER---VADPDTT 1069

Oy 1045 EKQABE-----EDEHLERTLEONKGMACKEE 1071

Db 1070 CESAREWMLSKFEQLDRNPNDVKSIGNLERDLKPLEDRALKLEK 1113

RESULT 11

SCHUB

spectrin beta chain - human

C:Species: Homo sapiens (man)

C:Date: 30-Jun-1992 #sequence revision 30-Jun-1992 #text change 09-Jul-2004

C:Accession: A37064; A39346; J70554; A28777; A39885; B23659; B42872; B27016; A32514

R:Winkelmann, J.C.; Chang, J.G.; Tee, W.T.; Coste, P.; Scarpa, A.L.; Marchesi, V.T.; Forget, B.G.

J: Biol. Chem. 265, 11827-11832, 1990

A:Title: Full-length sequence of the cDNA for human erythroid beta-spectrin.

A:Reference number: A37064; MUID:90307707; PMID:2195026

A:Accession: A37064

A:Molecule type: mRNA

A:Releases: 1-2137 <MIN>

A:Cross-references: UNIPROT:P11277; UNIPARC:UP10000135DB5; GB:J05500

A:Note: the authors translated the codon GAG for residue 596 as Arg, CAC for residue 801

R:Gallagher, P.G.; Tee, W.T.; Coste, P.; Scarpa, A.; Bolvin, P.; Delaunay, J.; Forget, B.G.

J: Biol. Chem. 266, 15154-15159, 1991

A:Title: A splice site mutation of the beta-spectrin gene causing exon skipping in heredit

A:Reference number: A39346; MUID:91332035; PMID:1840591

A:Accession: A39346

A:Molecule type: DNA

A:Releases: 2002-2137 <GAL>

A:Cross-references: UNIPARC:UP10000173DB5; GB:J05500

R:Ioan, S.H.; Kentros, C.G.; Pritch, J.T.

Gene 91, 297-302, 1990

A:Title: Identification of an unusual deletion within homologous repeats of human reticul

A:Reference number: J70554; MUID:91007291; PMID:1976574

A:Accession: J70554

A:Molecule type: mRNA

A:Releases: 928-957, 'NY', 960-982, 'V', 984-1030, 'D', 1032-1755 <YOO>

A:Cross-references: UNIPARC:UP10000173DB6; GB:M57948

R:Winkelmann, J.C.; Leto, T.L.; Watkins, P.C.; Eddy, R.; Shows, T.B.; Linnenbach, A.J.;

Blood 72, 328-334, 1988

A:Title: Molecular cloning of the cDNA for human erythrocyte beta-spectrin.

A:Reference number: A90733; MUID:88269838; PMID:3390609

A:Accession: A28777

A:Molecule type: mRNA

A:Releases: 1334-1373, 'R', 1375-1432, 1909-2137 <W12>

A:Cross-references: UNIPARC:UP1000016A607; UNIPARC:UP1000016A608; GB:J05500

A:Note: authors Yoon et al. show his to be the predominant residue at position 1374 and

R:Pitch, J.T.; Morley, B.U.; Yoon, S.H.; Coetzee, T.L.; Palek, J.; Conboy, J.G.; Kan, Y

Proc. Natl. Acad. Sci. U.S.A. 84, 7468-7472, 1987

A:Title: Isolation and characterization of cDNA clones for human erythrocyte beta-spect

A:Reference number: A39885; MUID:88041127; PMID:3478706

A:Accession: A39885

A:Molecule type: mRNA

A:Releases: 1209-1482 <PRC>

A:Cross-references: UNIPARC:UP1000016B05C; GB:M18054; NID:G338333; PIDN:AAA60572.1; PID:

A:Experimental source: reticulocyte

R:Winkelmann, J.C.; Costa, F.F.; Linzie, B.L.; Forget, B.G.

J: Biol. Chem. 265, 20448-20454, 1990

A:Title: Beta spectrin in human skeletal muscle. Tissue-specific differential processing

A:Reference number: A23659; MUID:91056094; PMID:2243099

A:Accession: B23659

A:Molecule type: mRNA

A:Releases: 2105-2137 <W13>

A:Cross-references: UNIPARC:UP10000173DB7; GB:M37885

R:Speicher, D.W.; Weglarz, U.; Desliva, T.M.

J: Biol. Chem. 267, 14775-14782, 1992

A:Title: Properties of human red cell spectrin heterodimer (side-to-side) assembly and i

A:Reference number: A42872; MUID:92340516; PMID:1634521

A:Accession: B42872

A:Molecule type: protein

A:Releases: 47-56/293-302, 1837-1846 <SP2>

A:Cross-references: UNIPARC:UP10000173DB8; UNIPARC:UP10000173DB9; UNIPARC:UP10000173DBA

R:Spelcher, D.W.; Marchesi, V.T.  
 Nature 311, 177-180, 1984  
 A>Title: Erythrocyte spectrin is comprised of many homologous triple helical segments.  
 A>Reference number: A93341; MUID:84295639; PMID:6472478  
 A>Accession: B27016  
 A>Molecule type: protein  
 A>Residues: 292-324, 'X', 326-329, 'Y', 331-332, 434-532; 718-734, 'V', 736-773, 'X', 775-777, 1036-1994-1997 <SP8>  
 A>Cross-references: UNIPARC:UPI0000173DBB; UNIPARC:UPI0000173DBC; UNIPARC:UPI0000173DBD;  
 A>Note: the purified protein had a blocked amino end  
 C>Comment: Spectrin is a major structural component of the erythrocyte membrane cytoskeleton  
 C>Genetics:  
 A>Gene: GDB:SPTR  
 A>Cross-references: GDB:119602; OMIM:182870  
 A>Map position: 14q23-14q33  
 C>Superfamily: spectrin beta chain; alpha-actinin actin-binding domain homology; spectrin  
 C>Keywords: actin binding; cytoskeleton; duplication; erythrocyte; heterodimer; membrane  
 F.153-271/Domain: alpha-actinin actin-binding domain homology <ACT>  
 F.301-412/Domain: spectrin/dystrophin repeat homology <SP01>  
 F.421-526/Domain: spectrin/dystrophin repeat homology <SP02>  
 F.527-635/Domain: spectrin/dystrophin repeat homology <SP03>  
 F.636-741/Domain: spectrin/dystrophin repeat homology <SP04>  
 F.742-846/Domain: spectrin/dystrophin repeat homology <SP05>  
 F.847-953/Domain: spectrin/dystrophin repeat homology <SP06>  
 F.953-1059/Domain: spectrin/dystrophin repeat homology <SP07>  
 F.1060-1166/Domain: spectrin/dystrophin repeat homology <SP08>  
 F.1167-1272/Domain: spectrin/dystrophin repeat homology <SP09>  
 F.1273-1377/Domain: spectrin/dystrophin repeat homology <SP10>  
 F.1378-1476/Domain: spectrin/dystrophin repeat homology <SP11>  
 F.1477-1582/Domain: spectrin/dystrophin repeat homology <SP12>  
 F.1583-1688/Domain: spectrin/dystrophin repeat homology <SP13>  
 F.1689-1795/Domain: spectrin/dystrophin repeat homology <SP14>  
 F.1796-1901/Domain: spectrin/dystrophin repeat homology <SP15>  
 F.1902-2007/Domain: spectrin/dystrophin repeat homology <SP16>  
 F.2008-2118/Domain: spectrin/dystrophin repeat homology <SP17>

```

Db      599 GYOPCDQVILIDRSMHLEQCEEEELSNMAAGARTQL-EQSKLMKRFEMMEDEAEEMIKKE 657
Qy      730 LPSYGEAAAK-----LKRSKASGDENDNINIDITNEIPE--GRVYGGSDLTMLN 780
Db      658 -QIYSSLDYGDLTSLIILQKHKAFEDERLGLDHLLEQIFQEAHGVNA-----RN 707
Qy      781 DLDP--EQNSKLVLDLKLLEVPQVANSPPSSAAQKAVTSSSEODMSGTEDLATERL 838
Db      708 EFGHPQIARIKVSAAQMDQLAALFCCKNLQDAENPFGQSGADDLKMLD----- 761
Qy      839 QKTERFRNPVPSKDSVTARKTQLOFSQYIENRPEMKRORSIQEDTKKNGEKAATET 898
Db      762 ---AHLRSGEDVQDQDEGATRALCKKMDPLE---ELIESRGVMEHLEQ-----QA 806
Qy      899 QRKPSDEVINKPKQKQSYVYVAGELALENQKQIDTRALVEKRLAYLMDTGNTSEEE 958
Db      807 QGPRPE-----FRD--SPDVTIRLQALRELYQVVAQADLRQORLOEALDLYTFGETD 858
Qy      959 AMQWEMLVNKKNALIRMANQLSLEKEHDERREYELNRELRAMLAIED----- 1009
Db      859 A-CELM---MGEKEMLAEMEMPTLEDLEVQCHRPDILOEMKTLTQIDGVLAANSL 914
Qy      1010 -----WQ--KTBAQRRERQL----- 1026
Db      915 VESGHPRSREVKQYQDHLNTRWQAFQTLVESRREAVDSALRVHTLCVDCSEETSKMTDKT 974
Qy      1027 -----LVVALNKKDALVRD-----LDAQEKQAE--EDHELTLEEQNK 1063
Db      975 KVESTQDQGRDLAGITAIQKLSGLERDVAAIQARVDALERESQQLMDSHPQ--KENT 1032
Qy      1064 GKMAKKEEK 1072
Db      1033 GQRQKHLE 1041

```

## RESULT 12

```

T29140
hypothetical protein K11C4.3 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C/Accession: T29140
R/Pauley, A.; Gaitung, S.
submitted to the EMBL Data Library, July 1996
A/Description: The sequence of C. elegans cosmid K11C4.
A/Reference number: Z20577
A/Accession: T29140
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-2326 <P>A>
A/Cross-references: UNIPARC:UPI000035888; EMBL:U64854; PIDD:AB18316.1; GSPDB:GN00023;
A/Experimental source: strain Bristol N2; clone K11C4
C/Genetics:
A/Genes: CESP.K11C4.3
A/Map position: 5
A/Intons: 105/1; 131/3; 155/3; 449/3; 2075/1; 2140/3; 2182/1; 2232/2; 2309/1
C/Superfamily: spectrin beta-G chain; alpha-actinin actin-binding domain homology; plect

```

Query Match 5.6%; Score 309; DB 2; Length 2326;

Best Local Similarity 20.1%; Pred. No. 2.2e-05;

Matches 201; Conservative 174; Mismatches 327; Indels 300; Gaps 44;

```

Qy      289 NASOSLLWCKEYKQYRGVKTNPETTSMRNGLSFCAILHPPDLIDYKSLNPQDIKEN 348
Db      229 SAREALLWQCMQKTAGPNNVNVKPFSTSRDGLAFNLIIKHHPDLVDYDNLQKSNALYN 288
Qy      349 NKRAYDGFAS-IGISRLLEPSDWLALIPDKLTWVTVLYQIRAFS-----GQELNV 399
Db      289 LQSAFTAEHQDLAKFLADE--VNVDQPBKSIITTVVYVYFNKDKDNIGKRIK 347
Qy      400 V--QIBNSKSYVYKGNVETDTNNSVD-----QEKYAEISDLKRPBLOQPI 447
Db      348 VINELMENDK---MINRVETLSSDLEWLNAKIQLINBRHFENNLEGVQRQ-----LT 397

```

```

Qy      448 GAVDFLSQDDSVFVNDGVESESEHQTPDDHLSPSTASPYC--RRTKSDTEPQKQSSG 506
Db      398 EFNVDYRQGEPRPFDER--GELEVLFTLQSAARANNQRFFVREGLIADIRANWSLE 455
Qy      507 RTSGSDDPGICSNITSDTAQVLLGKKRLKAETLE-LSDLVSDKCKKMSPPICE---- 561
Db      456 K-----AHERRELVKELIRQEKLEQLAAAF--NRKAMEMETWLTENQRL 499
Qy      562 -----ETD-----EQQLQ--TLIDGSNLEKKNLSLEERSD 593
Db      500 VSQDNFGNDLSVBAATYKKEALETDIPAYEEVQAVAAVAGELAEANYDQAKINERKE 559
Qy      594 PESPIKKTSLPTSKLGYSYRDLDAKKGAISIRQTESPPDADR--TLINAHDSKIVQ 652
Db      560 NVQL-----WNLIFQILLARKRRLSELMAIQIIFHDMILTLTLDMD-----IK 603
Qy      653 HRLISRQELKEBARVLLLEQARRDALKA-----GNKQNTNTA-----T 691
Db      604 SRLIS--EDLGAHMDVEDLLQKHALLSDINIIIGERVNNSIAQQRFRNPDGDSGYK 661
Qy      692 PFCGRQLSDQD-----ERRRQLABRARQL-----IAKSGVYKSE---- 729
Db      662 PVEPQTIDERSDVLQKRYKELDLAAERKRRLBDNKLCOFMDVVALEHGIKEQEVLS 721
Qy      730 -----LPSYGEAA-----EKLKERSKASGDENDNIEIDTNEIPEGV-- 768
Db      722 STDQGRDIYVSHLAKHKAENNLDEKYLDRLDVSGRELQDESIPEGSDNIPRLAEI 781
Qy      769 -----VGGDELTVLENDLTPB-----QNSKLVDLK 795
Db      782 RDYINKLKELSARKERLAGVEYQFFTDADVDRLVTLRWSESDVQDEGTVQL 841
Qy      796 LKQLLEVPQVAN-SPSSAAQKAVTSSSEODMSGTEDLATERLOQTERFRNPVPSKD 854
Db      842 LKQGDVHDELQFQDHIKVLAHAKASLPQEAR-HPDIR-QRLDTTLK----- 888
Qy      855 STVFKTQLOFSQYIENRPEMKRORSIQ-----EDTKKNGEKAATET---QR 900
Db      889 ---QKAELENTSG-----LRKQRLIDALSILYKLYSDADSVSMIDKCKKLATLVGR 938
Qy      901 KPSEDEVINKGFKDTSQYVVGELALEN-----EQKQIDTRAL 939
Db      939 DIREVEIMKRFPTLBDQMNQAKVYTNVDLARQLLNVEHPNSDDILHQNKLARWA-- 997
Qy      940 VEKRLAYLMDTGNTTEEBEEMQEMFML-----VNKNALIR---RMNQLSLEKE 987
Db      998 ---QLRDVNDQKKN-ELERAHRLLETRIDQETVTWIEBKTRVLESDALTNDLGVMK- 1052
Qy      988 HDLERREYELNRELRAMLA-----IEDMOKTBAQRRERQL-LDELVALVNRKD 1035
Db      1053 --LQRKLSMWERDLGAIQAKLDSLHKAADDIERERQEAQAIEDIKRIHQVWDILANK- 1109
Qy      1036 ALVRDLDQEKQAE-----EDHELTLEEQNKMAKKEE 1071
Db      1110 --VRHBKAKLDEAGDQLQRFRLRDLDHQAULTATQROVASSEE 1149

```

## RESULT 13

T18296

myosin heavy chain - Entamoeba histolytica

C/Species: Entamoeba histolytica

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T18296

R/Guillen, N.

submitted to the EMBL Data Library, February 1997

A/Reference number: Z18865

A/Accession: T18296

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-2139 <G>U>

A/Cross-references: UNIPROT:Q07569; UNIPARC:UPI000080203; EMBL:L03534; NID:G1850912; PIR

A:Gene: mhca  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
F:91-780/Domain: myosin motor domain homology <M0>

Query Match 5.4%; Score 299; DB 2; Length 2139;

Best Local Similarity 20.2%; Pred. No. 5.2e-05;  
Matches 254; Conservative 213; Mismatches 448; Indels 340; Gaps 53;

```

QY 38 FLREGKATDEDMQST-----ASLVSKQKQDIGNLDPEEDNEDDENRNVQ 83
DB 859 FBEERKQKDEKEDLKKGLAEIKKGRBAENALASATKATGELAKIQDLED----- 910
QY 84 EKAAKITELINKLNF--LDEAKDLATVNSNPDPDPAEL-----NPFGD----- 128
DB 911 -----KISELSKLSAAELDKQELNLKIENTL-----EEDKSELKETIDMLKQDLKQSKLG 961
QY 129 PDSEPIETETASPKRTIESFYNNSNPKREVOTPOYLNPFPPEAFVT-----I 177
DB 962 EDLEVEITELNSQIMTLNATVDKDKDTIAEMQ-----ESIDEKEDEITKLGKIDILBEEB 1016
QY 178 KDSPPQSTKRNIRPVDMSKTYA-----DSKTEBELEDSNPFYEPKSTPPPN 227
DB 1017 KQDLBQDRADVATKQDIAKLNKTTICEBQKQDIATLBOLEDEEN-----KXK 1067
QY 228 NLVNPVQEL-----ETERRVYKAKAPAPVLSPTGVLNENTVSAQKDLSTPKSPISP 282
DB 1068 DLTEBLQOTQKLGSTEKSLAAQVAATKASDERLTLSQN----- 1107
QY 283 VLGRKPNASQGLVWCKEVTYKNGVKITNFTTNSRNLSPCALIHHFRPDLIDYKSLNP 342
DB 1108 -LENEKLTAKMLTKADLEKKSISGLK-----QDYEDL-- 1139
QY 343 QDINENKKAADGSPASIGISRLBESDNLVLAIPDKLT-----VMTLYQIRAHNSGQ--- 395
DB 1140 -----BDDKNKTEGD-----LNAQKRIKLELDITKQADVQYLOKQKEYESQIAK 1187
QY 396 ---ELNVQVQIENSSEKSTYKGVNGETDT--NSSVDQEKYAEISDKREPELQOPISGAVD 451
DB 1188 MQEKEKAIQNVKVKKEKTIKEKELEIOGLOEKLBTEVYKDAE--KKGKELEKEMKAL-- 1244
QY 452 FLSGDDSVFVNDGVSSESEHQTPPDHLSPS-----TASPYCRRTYS--DTEPQSKSQ 503
DB 1245 ---OEKKNV--ESSKNSTEEKKKLEKLDNLIKDTQKKLDQMTADNEKLLKAKAKDLQALBKV 1300
QY 504 SSGRTSGSDPDGICGNTDSTQAVLIGKKRLIAKTELESLDLYSDKQKQMSPPICGET 563
DB 1301 QDNHKAVALAEELNKKKAQSDKEKLSLKAELEALTKAKSVESKQKQSENEKALSEBI 1360
QY 564 DEQKLQTLDIGSNLEKEKLENSRSLSECRSDPESPICK---TSLSPSTKLGYSGRDLD-- 618
DB 1361 DQANEKIKNIQADLRKATADLQEANKEKAYEAQDKLVADNKKMTTLEIKAKADEENT 1420
QY 619 -----LAKKQASLRLQTSDDPDADRTTNHADHSKIVQHLLSRQEBLK-----ERA 666
DB 1421 YKVENYEVYLKRKEADLEBANENLDIEKQDNMKKEQYKQKGELEKTKDNLMAIAEKD 1480
QY 667 RVLLEQARDAALAKGK---HTNTATTPCNRQLS-----DQDERRRQALERRA--- 714
DB 1481 SIFPAKQSDADLEBLNKTVEHEHDEVAK--LNTQITLTLDNDQSAEBELNELSKDKD 1538
QY 715 ROLIAEARSQVGMELPSYGEMLAEK-----LKRSRASGDENINIBDNTN- 761
DB 1539 KKKISELEBOVNEBSRPGVTGNADENEIKRDAQIADLNLKALEKQVQNNQLOQTNKEI 1598
QY 762 -----EIPGFWVGAGDBLTNLENDLDTPEQNS-----KLVLDTKL 796
DB 1599 KAKONDLSKIEITE-----NEKKLENAKKQLEQDKDADKAVSQGTIKRGLSEEV 1651
QY 797 KKL-LFVQP---QVANSFSSAAQKAVTSSSDQKMSGTFEDLRTERL-----QKTER 844
DB 1652 KKLTFEIOALKFQI--NAPSSVAOEBEKKQRLSDIALKLBQLEOERTTAANAABARKKIQ 1710
QY 845 FRNPFVPS-KDSTVRKQTQLOSPFSQYIENRPFMKQKQRSIQEDTKKGNNEKAITERTQKPS 903

```

```

DB 1711 ELDEVKFNLBDVTNQREKL--VAKNSENDADID---SLKEKKKALDEDEIKTIDDNNGLS 1765
QY 904 E-----DEVIAKGFQDTSQYVVGELA-----ALENEQK----- 931
DB 1766 EIDISLQKRYALLDSKQSDVSMKEKQD-----BLKTKQDLLEFEKKNHAETMRLLKG 1818
QY 932 QIDTRALVREKRLAYL-----MDTGRNTESEBAMMQEWFMLVNNKQALIRRN--QL 981
DB 1819 RLKEKAAEVQRLNLAQKNDLDAQQEKAKATKDRADADGELKSLNLELDVYQDLDRKQD 1878
QY 982 SLLEKHD--LERRYELNR-----ELRAMLAIEDQKTEAQKRE--QLLLDE 1026
DB 1879 DLADKEDLATLQDKYKTLVQKSVPSRIGEMQEQDLDEKAGRAKQKQKQAYEKKIQE 1938
QY 1027 LVALVN-----KRDALVRDLDAQEKAESDEHLERTLEQNGKRAKK--EKK 1073
DB 1939 LQENDNDFEYKETADRKINTLSAQKDLQKELEK--ERGLQDSEKEVQRLRVKX 1992

```

# RESULT 14

```

528916
dystrophin - mouse
N:Alternate names: duchenne muscular dystrophy protein
C:Species: Mus musculus (house mouse)
C:Date: 22-Nov-1993 #sequence revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S28916; B27162; S10922; C43837; B40134
R:Bies, R.D.; Phelps, S.F.; Cortez, W.D.; Roberts, R.; Caskey, C.T.; Chamberlain, J.S.
Nucleic Acids Res. 20, 1725-1731, 1992
A:Title: Human and murine dystrophin mRNA transcripts are differentially expressed during
A:Reference number: S28916; MUID:92253376; PMID:1579466
A:Accession: S28916
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-3678 <BIE>
A:Cross-references: UNIPROT:P11531; UNIPARC:UPI0000279E7; EMBL:M68859
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991
R:Koenig, M.; Hoffman, E.P.; Bertelson, C.J.; Monaco, A.P.; Feener, C.; Kunkel, L.M.
Cell 50, 509-517, 1987
A:Title: Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary
A:Reference number: A90897; MUID:8727512; PMID:3607877
A:Accession: B27162
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-201 <KOB>
A:Cross-references: UNIPARC:UPI0000177686
R:Mandel, U.; Zuk, D.; Elnat, P.; Zeelon, E.; Levy, Z.; Neuman, S.; Yaffe, D.
Nature 337, 76-78, 1989
A:Title: Duchenne muscular dystrophy gene product is not identical in muscle and brain.
A:Reference number: S06461; MUID:89082658; PMID:2909892
A:Accession: S10922
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-106 <NUD>
A:Cross-references: UNIPARC:UPI0000177687; EMBL:X14183
R:Rapaport, D.; Lederfein, D.; den Dunnen, J.T.; Grootscholten, P.M.; Van Ommen, G.J.; R.
Differentiation 49, 187-193, 1992
A:Title: Characterization and cell type distribution of a novel, major transcript of the
A:Reference number: A43837; MUID:92316332; PMID:1377655
A:Accession: C43837
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 3069-3181 <RAP>
A:Cross-references: UNIPARC:UPI0000177688
A:Note: sequence extracted from NCBI backbone
R:Hoffman, E.P.; Monaco, A.P.; Feener, C.C.; Kunkel, L.M.
Science 238, 347-350, 1987
A:Title: Conservation of the Duchenne muscular dystrophy gene in mice and humans.
A:Reference number: A40134; MUID:88018015; PMID:3659917
A:Accession: B40134
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 300-676, 'F', 678-1390 <HOF>

```



```

Db      1446 ARDAVEHLTKSLA-----DVEQVSVQNO-----EKDALGRL 1478
Qy      533 RLKAEITELSDLYVSDKKOMSPFICEPDEQLQTLIDISNLE-----KE 580
Db      1479 ALLORE-----RDK-----LIVEMDKSLENOISLGGSCSLKALGLTDEK 1521
Qy      581 KL-ENSRSLGCSDESPDKTSLSPTSK-----LGYSYRDLDAKK----- 622
Db      1522 KLMKELESVRCSKIAESTEMQEKHELOREYEVLLQSTENNSNEARIQHVYESVROEKQ 1581
Qy      623 -KHASLRQTESDPPDADRITLHNDHSSKIYQHRLLSROBLKERAR-----VLBOA 673
Db      1582 EYVALRRAESDPRERERKQLOAE-----QEMEMEKERKRFKSKQKILELE 1630
Qy      674 RRDALKA-----GNKHTNTATPFCNRQLSDQDE---ERRQRLR-----RAQLIAE 720
Db      1631 BENDRLRAEAPVGGANESMEALSSNASLKEBELERTLEYKTLSEKEFEALMAERNTLSE 1690
Qy      721 ARSGVQNSELPSYGEMAEKLERESKASGDENDNIETNEBIEPFGVVGDELTNLEN 780
Db      1691 ETRNKL-----QVEAQELKQASLETTEKSDEPK-DVIEVYEA-VVGKSQE----- 1735
Qy      781 DLDTPQNSKLVDLKLLKLEVPQVANSPPSAAQAVTE--SSBQDKSGTEBL----- 833
Db      1736 -QDSISENAKLEDAEATLL-----ANS---AKPGVSETPSSHDIINNYLQQLDQLKG 1783
Qy      834 -----RTERLOKTERFRNPV---VPSKOSTVR---KTQLOSFOYIENRPEM 875
Db      1784 RIABIEMEKQORELSOTLENEKNALLTOISAKOSELKLEBEVAKINMLNQIOE--EL 1841
Qy      876 KRQSIQEDTKKGNBE-----KAAITETQKPSDEVLNKGFKDTSOVVGEI 923
Db      1842 SRVTKKETAEERKDDLEERLMOALBNLSIGNTYQVDTAQIKNQOLESBMQNLKRCV 1901
Qy      924 AALENEQKQIDTRALVYKRLR--YL-----MDTGRNTEBBBAMQEWPMVLVNKNAL-- 974
Db      1902 SELEBEKQOLVBKTKVSEBIRKEVMEKIQGAQKGPSSKIHAKELQE--LLKEKQOEVKQ 1959
Qy      975 -----IRRMQLSLEKHEHDLERYEELNRELRLML-----AIEDMQTEAQKRE 1020
Db      1960 LQKDCIRYLRISALEK---TVKALFVHTESOKDLATKGNLAQAVEHHKKAQAEISSP 2016
Qy      1021 QLLDDE-----LVALLVNRDALVPRDLDAQEKQAEEDHLEERTLEFONKGMAKKEEK 1072
Db      2017 KILHDDTQSEARVLAADNLKLELOSNKESIKSQIKQKDDLLRRLQAEK-HRREKK 2075

```

Search completed: May 20, 2006, 17:15:34  
 Job time : 59 secs

**THIS PAGE BLANK (USPTO)**

This Page is inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLORED OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images  
problems checked, please do not report the  
problems to the IFW Image Problem Mailbox**



**THIS PAGE IS BLANK**